

METHOD IN MIND

A METHODOLOGICAL PERSPECTIVE ON THE CHILD-AS-SCIENTIST DEBATE

A thesis
submitted in fulfilment
of the requirements for the Degree
of
Doctor of Philosophy in Psychology
in the
University of Canterbury
by
Claire F. O'Loughlin

University of Canterbury
2002

Acknowledgements

Sincere thanks firstly to my supervisor, Brian Haig, for his encouragement throughout the course of this project, and for his flexible and creative approach to teaching and supervision more generally. Thanks are also due to Paul Thagard for a stimulating 10-week research fellowship to his cognitive science laboratory at the University of Waterloo.

I owe an enormous debt to family and friends for their support over the years and, in particular, to my parents for housing and feeding me in the final stages of writing.

Finally, this thesis is dedicated to Alex, for enduring long absences, late night phone calls, endless reading of chapters, and much more.

Table of Contents

List of Tables	viii
List of Figures	ix
Abstract	x

Introduction

Children's knowledge and framework theories	3
Knowledge systems and strong restructuring	4
The child as spontaneous theoretician	5
Misrepresenting children and science	6
What constitutes a productive scientific analogy?	8
Basic principles	9
The multiconstraint theory of analogy	12
A template	16

Chapter 1 The Child-as-Scientist Analogy: Current Research and Applications

1.1 The Theory Theory	19
1.1.1 Science and childhood: The cognitive connection	21
1.1.2 Theories and theory change	22
1.1.3 Alternative accounts	24
1.1.4 Empirical defence	25
1.2 Questioning the extant framework	26
1.2.1 Children and scientists: Analogy or identity?	26
1.2.2 Modularity, maturation, and baby theorists	29
1.2.3 Kuhnian revolutions and cognitive mechanisms of change	31

1.3	Refocusing the debate: A methodological perspective	35
------------	--	-----------

Chapter 2 Choosing a Source Model

2.1	Science as method	38
2.1.1	Evolutionary epistemology: An overview	38
2.1.2	Implications for understanding science	41
2.2	Theory-evidence relations: The key to explaining cognitive development	43
2.2.1	Revisiting Socrates' and Augustine's problems	43
2.2.2	Theories and evidence: How does knowledge develop?	45
2.3	Orthodox theories of scientific method	48
2.3.1	An inductive account	48
2.3.2	A hypothetico-deductive account	52
2.4	Beyond orthodoxy	55

Chapter 3 An Abductive Theory of Scientific Method: A Third Alternative

3.1	Detecting and explaining empirical phenomena: Twin goals for inquiry	58
3.1.1	(Re)focusing on empirical phenomena	58
3.1.2	The role of abduction in scientific explanation	66
3.1.3	Abductive method: An overview	72
3.2	The abductive framework	75
3.2.1	Research problems	75
3.2.2	Phenomena detection	76

3.2.3	Theory generation	79
3.2.4	Theory development	81
3.2.5	Theory evaluation	82
3.3	The abductive theory as a general theory of scientific method	84

Chapter 4 From Source Model to Target: Mapping Relations Between Scientists and Children

4.1	“Interactions all the way down”: An alternative perspective on development	89
4.1.1	The story so far	89
4.1.2	Navigating between the poles of nativism and empiricism: Attempts to chart a viable “radical middle”	91
4.1.3	Development as a truly interactive process	94
4.2	Theory building in scientists and children	102
4.2.1	Narrowing the focus: Repositioning the child-as-scientist analogy	102
4.2.2	The child as a spontaneous theory builder	103
4.2.3	Detecting and explaining empirical phenomena: A framework for investigating theory building across scientific and everyday contexts	109
4.3	Beyond the theory theory	113

Chapter 5 Evaluating Ideas: The Role of Coherence in Scientific and Everyday Thought

5.1	Applying theories in scientific and everyday contexts	116
5.1.1	Children’s theories of the natural world: Examples from observational astronomy	117

5.1.2	Evaluating explanations: Criteria for theory choice	126
5.2	Theory building, evaluation, and explanatory coherence	130
5.2.1	The theory of explanatory coherence (TEC)	131
5.2.2	Computing explanatory coherence: An introduction to ECHO	134
5.2.3	Do children apply principles of explanatory coherence?	137
5.3	The pervasiveness of coherence-based inference	143
5.4	Appendix	145

Chapter 6 Speculating on the Cognitive Origins of Science

6.1	The theory theory's solution to an evolutionary puzzle	147
6.1.1	Science is a 'by-product' of childhood	149
6.1.2	Evolved mechanisms and scientific progress	151
6.2	What makes scientific inquiry possible?	154
6.2.1	Have Gopnik and Meltzoff solved the evolutionary puzzle?	154
6.2.2	The emergence of scientific abilities in human evolution	159
6.3	Abduction, tracking, and science	164

Chapter 7 An Abductive-Methods Perspective on the Child-as-Scientist Debate

7.1	Looking back	168
7.2	Comparing analogies: The theory theory versus the abductive-methods account	172
7.2.1	The theory theory	172

7.2.2 The abductive-methods account	178
7.3 Future directions: Consolidating an abductive perspective on human reasoning	182
References	189

List of Tables

Tables	Page
I.1 The child-as-scientist analogy: Examples of attribute, relational, and system mappings	16
6.1 The emergence of key elements of science in human evolution	161
7.1 A representation of the theory theory using the multiconstraint theory's taxonomy of attribute, relational, and system mappings	173
7.2 A representation of the abductive-methods account using the multiconstraint theory's taxonomy of attribute, relational, and system mappings	178

List of Figures

Figures	Page
2.1 The main inferential move in an inductive account of scientific method	50
2.2 Hypothetico-deductive method	53
3.1 Curd's (1980) classification of the components of a logic of discovery	69
3.2 The explanatory move from phenomena to theory in the abductive method	73
4.1 Claims on the "radical middle" position: Attempts to define a new theory of cognitive development at the 'intersection' of nativism, empiricism, and constructivism	92
5.1 Children's mental models of the earth	119
5.2 Drawing of the earth, moon, the stars, and the sky by Venica, Grade 3 (hollow sphere)	121
5.3 Example of a mental model of the day/night cycle constructed by children in Vosniadou and Brewer's study that is simpler than the scientific model	125
5.4 Experimental materials used to test children's application of the range of explanation criterion and the non-ad hocness criterion	127
5.5 A simple ECHO network	136
5.6 An ECHO network depicting the 'range of explanation' condition on the chemistry task	140
5.7 An ECHO network depicting the 'non-ad hocness' condition on the chemistry task	141

Abstract

During the past decade, an extensive body of cognitive developmental research has emerged that subscribes either implicitly or explicitly to the analogy of the child as an intuitive scientist. This analogy captures the idea that children resemble scientists in their attempts to explain and predict phenomena, and thereby directs research attention to the development of knowledge in science as a potential source model for cognitive development. Foremost among proponents of this analogy are Alison Gopnik and Andrew Meltzoff, whose theory theory account of child-scientist parallels takes the processes subserving theory change in science and childhood cognitive development to be essentially the same. This thesis critically examines the theory theory and proposes an alternative formulation of the analogy in which the construction of meaningful relations between children and scientists is achieved by developing a methodological perspective on the debate. With this aim in mind, the Introduction comprises a brief review of the recent history of child-scientist parallels in cognitive developmental research and offers some methodological provisions for constructing a productive scientific analogy. Chapter 1 then undertakes a detailed examination of the theory theory, and argues that the major limitation of Gopnik and Meltzoff's account lies in its continued reliance on an inappropriate source model. Following this critique, Chapter 2 presents a general argument for a methods-centred model of child-scientist parallels and examines two orthodox theories of scientific method as candidate source models for the analogy. Having argued that neither account has the capacity to provide an adequate account of scientific inquiry, Chapter 3 details a comprehensive abductive theory of scientific method, and argues for its adoption in a methodological reformulation of the analogy. With this source model in place, Chapter 4 aims to establish meaningful relations between the abductive account of scientific method and the knowledge construction efforts of young children by defining a narrower role for the child-as-scientist analogy within an interactionist account of development. Working within this framework, compelling parallels are identified between the theory building strategies of scientists highlighted by abductive method and the data-to-theory moves uncovered in microgenetic analyses of children's problem solving. Having identified an abductive pattern of reasoning in children's spontaneous theory building, Chapter 5 extends the analogy to the processes by which children evaluate the quality of everyday explanations, drawing

on a computational model of theory evaluation for this purpose. In Chapter 6, the focus turns to speculations regarding the cognitive origins of science, in light of the theory theory's appeal to an evolutionary warrant for child-scientist parallels. Finally, Chapter 7 provides a summary of the main arguments and reinforces proposals for the utility of a methodological perspective on the child-as-scientist debate by undertaking a detailed comparative evaluation of the theory theory and the abductive-methods account. A discussion of future directions concludes this study.

“To propose an analogy, or simply to understand one, requires taking a kind of mental leap. Like a spark that jumps across a gap, an idea from the source analog is carried over to the target. The two analogs may initially seem unrelated, but the act of making an analogy creates new connections between them. Nothing ever guarantees that the target will actually behave the way the source suggests it should . . . Some of the mental leaps accomplished by analogy have ended in creative triumphs; others have ended in dismal failures. Analogy must be recognized as a source of plausible conjectures, not irrefutable conclusions. The success of an analogy must finally be judged by whether the conjectures it suggests about the target analog prove accurate and useful.”

(Holyoak & Thagard, 1995, p.7)

“Often, progress in science begins with finding the right analogy. Recently, cognitive and developmental psychologists have invoked the analogy of science itself.”

(Gopnik, 1996b, p.485)

Introduction

How can we best characterize the development of knowledge in childhood? Do young children know different things from adults about the world, or is their knowledge essentially the same as ours? Are they immature thinkers, or do the forms of reasoning seen in childhood tell against any strong demarcation line between early and later cognition? What about the origins of knowledge? What representations and/or cognitive structures do we begin with and what particular forms do they take; for example, is it informative to characterize these structures as modules, constraints, biases, or 'starting state' theories? Further, does the current picture we have of the developmental process suggest an orderly sequence of incremental shifts, or are there major reorganizations of knowledge occurring? Finally, what role does the child's everyday experience play in this progression? Does it simply act as a catalyst for maturing modules or more actively function as evidence for or against children's developing representations of the world? Or is the responsibility for development even more equally weighted between a structured environment and predispositions in the individual, with cognitive and linguistic abilities somehow emerging as the 'end products' of a dynamic interactive process?

The above questions highlight issues at the heart of cognitive developmental research. The content of children's knowledge, their reasoning processes, an accurate characterization of the initial state, the nature of conceptual change, and the role of experience, are all central to understanding cognitive development because they contribute to the task of describing and explaining the changes that occur in the developing mind. However, these questions are also seen to be relevant because they are informed by an increasingly dominant perspective on cognitive development that draws insights from the development of knowledge in science.

During the past decade, an extensive body of cognitive developmental research has emerged that subscribes either implicitly or explicitly to the analogy of the child as an intuitive scientist. This analogy captures the idea that children resemble scientists in their attempts to explain and predict phenomena, and thereby directs research attention to the development of knowledge in science as a potential source model for understanding cognitive development. The “theory theory”,¹ as this perspective has become known, receives its most detailed treatment in a recent book by Alison Gopnik and Andrew Meltzoff (1997).

In *Words, Thoughts, and Theories* (1997), Gopnik and Meltzoff claim that a single model of development applies across science and childhood. They argue that children’s and scientists’ conceptions of the world constitute theories, that conceptual development in both domains amounts to an ongoing process of theory construction and revision, and that semantic development is intimately tied to this theory building process. According to this version of the theory theory, we begin life with a number of “starting state theories” which are open to revision, and the subsequent learning process is largely one of theory change based on experience. Crucial to differentiating this view of cognitive development from alternatives that posit innate modules or the maturation of information processing capacities, is the claim that the nature of the relations between evidential input and theories is similar in both childhood and in science. Hence the task of capturing the developmental process in both domains, according to Gopnik and Meltzoff, involves specifying the causal relations that exist between initial theories, evidential input, and new theories that result from a conceptual restructuring process.

In this thesis, I critically examine the utility of the theory theory as presented by Gopnik and Meltzoff and propose an alternative formulation of the analogy in which the construction of meaningful relations between children and scientists is achieved

¹ This phrase was originally introduced by Adam Morton (Morton, 1980) to refer to the claim that our commonsense psychological understanding or folk psychology has the structure of a genuine empirical theory, a claim that he argued to be false. Advocates of the theory formation view, however, adopted the phrase and a developmental version of the theory theory has come to dominate the literature examining the emergence of an understanding of mind in childhood. More recently, researchers such as Alison Gopnik and Andrew Meltzoff have been employing the phrase far more widely to advocate a general ‘theory formation and change’ perspective on cognitive development. It is in the current broader sense that the phrase is used in the current work.

by developing a methodological perspective on the debate. I will argue that such a reinterpretation of the child-as-scientist analogy avoids the problems encountered by current attempts to construct mappings between children and scientists, and affords researchers a better platform from which to investigate the development of knowledge in childhood. With this aim in mind, it is useful to begin by situating the current presentation of the theory within a recent historical context that surveys a range of positions taken by researchers on the question of child-scientist parallels.

Children's knowledge and framework theories

An influential application of the scientific analogy to the content and structure of children's commonsense knowledge is found in the work of Henry Wellman (Wellman, 1990; Wellman & Gelman, 1992). In a seminal book investigating the young child's developing understanding of the mind (Wellman, 1990), Wellman puts forward an extended argument for the theoretical status of such knowledge. He argues that like scientific theories, children's early understanding of the mind can be characterized as coherent, as incorporating fundamental ontological distinctions or commitments, and as providing a causal-explanatory scheme for understanding phenomena in its domain. More particularly, Wellman proposes that the child's theory of mind takes the form of a framework theory and suggests that it is at this level of analysis, rather than at the level of specific theories, that the analogy with scientific knowledge structures is most informative.

Drawing on the work of philosophers who have attempted to capture the role of theoretical frameworks in science (e.g., Kuhn, 1962; Lakatos, 1970; Laudan, 1977), Wellman argues that a number of features of framework theories make them a useful source model for attempts to characterize the nature of our everyday conceptual knowledge. Firstly, framework theories serve to define what exists in a domain and can be seen to act as framing devices that constrain and support more specific attempts to theorize about phenomena. Secondly, these framework theories are relatively protected from empirical test and come under scrutiny only when a viable alternative causal-explanatory framework is made available that offers a replacement for the existing framework. Thirdly, framework theories are particularly relevant to issues of how knowledge develops since they constrain alterations to specific

knowledge structures within their scope and offer a way of thinking about large-scale change within a particular domain (Wellman, 1990). According to Wellman, the idea that our everyday knowledge inheres in such framework-theoretical structures allows us to make sense of the content-dependent, domain-specific nature of human cognition, and offers the beginnings of a research programme for investigating the development of these knowledge frameworks in childhood.

Knowledge systems and strong restructuring

While the intent of Wellman's work has been to achieve a more precise notion of everyday theories by drawing on philosophical accounts of framework theories in science, Susan Carey (1985, 1992) has looked to apply the analogy to the issue of how knowledge is reorganized during development. More specifically, she has investigated the possibility of substantive structural correspondences between knowledge systems in childhood and science and in particular, whether changes observed in children's conceptual systems involve strong restructuring of the sort seen during periods of revolutionary theory change in science (e.g., Kuhn, 1970).

In her 1985 case study of the emergence of biological knowledge (Carey, 1985), Carey attempts to answer these questions by undertaking a close examination of children's knowledge of animals and living things, and tracking the changes that occur in their understanding between the ages of 4 and 10 years. Based on her findings, she endorses the view of intuitive knowledge as organized within domain-specific theoretical systems and suggests that the reorganization of such knowledge over the course of development is best understood in terms of belief revision and conceptual change. More particularly, by highlighting evidence of a shift from a folk psychological framework and its attendant explanatory mechanisms to a genuinely biological framework, together with changes at the level of individual concepts, Carey argues that the extent of reorganization involved exhibits parallels with conceptual revolutions in science.

For Carey, then, the analogy with science demonstrates utility because it facilitates the construction of a more accurate description of knowledge acquisition in childhood. Specifically, it draws attention to the organization of children's knowledge within

conceptual systems, and reclassifies children's early knowledge as alternative models of the world rather than incorrect or impoverished versions of adult theories. In addition, the analogy promotes a concern with changes to the structure of these knowledge systems over the course of development and provides developmentalists with some much needed tools, such as Kuhn's doctrine of incommensurability (e.g., Kuhn, 1982), for tackling the problem of conceptual change.

The child as spontaneous theoretician

A somewhat different line of developmental research that makes use of child-scientist comparisons has been undertaken by Annette Karmiloff-Smith (e.g., Karmiloff-Smith & Inhelder, 1974; Karmiloff-Smith, 1988). In an article entitled *The child is a theoretician not an inductivist* (1988), Karmiloff-Smith focuses on the cognitive processes involved in discovery and describes a range of problem solving tasks designed to uncover the strategies young children use in their investigations of the physical world. From a microgenetic analysis of children's performance on these tasks, Karmiloff-Smith identifies a robust pattern of recurrent interconnected phases, indicating that children are engaged in spontaneous theory construction.

During the initial phase of discovery, children are "data driven" and their problem solving success rests primarily on the utilization of information in the external environment. Having succeeded on the task, children become "theory driven", suspending their interest in achieving behavioural success in favour of generating a rudimentary explanatory theory of the phenomena involved. Following this "internally driven phase", data and theory are realigned and the children once more achieve success on the task, but this time via the application of a consolidated and generalized theoretical framework. Based on her findings, Karmiloff-Smith concludes that children cannot be accurately characterized as inductivists; rather, like scientists, theory building constitutes a pervasive feature of their everyday problem solving.

From this brief summary of Karmiloff-Smith's research, it is obvious that the scientific analogy functions in a very different capacity compared to its role in Carey's investigations. Instead of considering structural parallels between conceptual

change in children and theory change in science, Karmiloff-Smith endorses the analogy primarily as a way of teasing apart children's action sequences from the internal mental representations underlying these sequences. By granting children "theories-in-action", and by focusing on the dynamic interplay between theory building, heuristics and data in children's problem solving, Karmiloff-Smith is able to explain a number of important cognitive changes that occur across development, in particular the U-shaped patterns of behavioural success on many tasks. More generally, her synchronic approach to the question of how knowledge develops, indicates that the processes of scientific knowledge construction are not qualitatively distinct from those found to be operating in everyday problem solving, and that theory construction plays a pivotal role in learning about the world in both contexts.

Misrepresenting children and science

In addition to those who endorse the analogy, a number of researchers have argued that comparisons between children and scientists promote misleading assumptions about both the nature of cognitive development and the nature of science. These arguments can be characterized as taking one of three forms: empirically based arguments concerning the reasoning processes employed by children, lay adults, and scientists (Kuhn, 1989); arguments from the general standpoint of developmental theory (Russell, 1992); and arguments based on social-constructionist analyses of the history of science (Gellatly, 1997). I briefly consider each in turn.

Based on empirical investigations of scientific reasoning skills in children and lay adults, Deanna Kuhn (Kuhn, 1989; Kuhn, Amsel & O'Loughlin, 1988) argues that the child-as-scientist analogy promotes an inaccurate view of commonsense reasoning. In support of this claim, Kuhn begins by defining a view of scientific thinking that is centrally concerned with the co-ordination of theories and evidence. More specifically, she proposes that the abilities required to explicitly state one's theory, to know what sorts of evidence would support it and what sorts of evidence would undermine it, and to justify its acceptance over competing theories on the basis of this theory-evidence co-ordination process, comprise the core skills involved in scientific reasoning. Following this definition of scientific inquiry, Kuhn reviews a range of research findings to suggest that these scientific skills undergo "strong restructuring"

over the course of development. She concludes that “the ability to reason scientifically” is not an accurate description of the knowledge seeking endeavours of young children. Rather, scientific thinking processes are best understood within a developmental framework in which the progressive differentiation and co-ordination of theories and evidence is seen to develop in tandem with the emergence of metacognitive capacities.

Another criticism of child-scientist parallels questions its utility as an explanatory framework for developmental inquiry. James Russell (Russell, 1992) argues that theory theorists’ accounts of the child’s emerging understanding of mind dangle at a hopelessly elevated level of description and, as a result, actively inhibit any useful discussion of the developmental process. In his view, focusing on theories and theory change is unhelpful for the following reasons. Firstly, such a focus neglects the cognitive abilities and competencies necessary for acquiring an integrated body of knowledge such as a concept of mind. Secondly, it offers no insights about the acquisition process itself, merely describing development as a succession of theories. Thirdly, a concern with theory formation and change fails to demonstrate any connective threads with the lower level of information processing and (particularly critical for Russell) executive control. Russell takes these limitations as motivation for dispensing with discussions of theories and theory change altogether, rejecting the analogy with science based on its failure to demonstrate any explanatory force.

Russell’s absence-of-mechanism argument is reinforced and further elaborated by Angus Gellatly (Gellatly, 1997). In an extended critique of the analogy, Gellatly works backward from the history of science to highlight problems of equivalence that arise when institutional change is adopted as a model for individual theory construction. He argues that comparing conceptual change in the young child with theoretical developments in the history of science promotes a fundamental “category error”. That is, it conflates the cognitive development of an individual with the social processes identified by philosophers (e.g., Kuhn, 1962) to be responsible for scientific theory change. Failing to respect this distinction between “the personal and the social” has, according to Gellatly, two unwelcome consequences. Firstly, it encourages a misrepresentation of scientific practice that masks its inherently social

nature. Secondly, it leads developmental researchers astray in their search for plausible mechanisms of change that are responsible for cognitive development.

In the preceding pages, I have sketched a recent historical context for the theory theory by highlighting the ways in which researchers have variously constructed mappings between children and scientists and pointed to some of the criticisms that have been levelled at these attempts to reason by analogy. In the chapters that follow, I intend to propose and develop an alternative model of the child-as-scientist analogy that situates the question of parallels between children and scientists at the methodological level. First, however, I suggest that attempts to achieve a clearer model of child-scientist parallels will benefit from a prior understanding of 1) the relations involved in analogical reasoning itself; and 2) the ways in which an effective analogy can advance scientific knowledge in a particular domain. Given that the overall aim of this thesis is to exploit the analogy between children and scientists in order to develop a more informative model of cognitive change, I therefore begin this model development process by examining the nature of analogy and its role in scientific inquiry.

What constitutes a productive scientific analogy?

A review of the literature on models and the use of analogy in science reveals that, traditionally, philosophers have been reluctant to credit modelling practices with a legitimate methodological role.² For example, Pierre Duhem, commenting on the mechanical models employed by 19th century English physicists suggests that, at best, such devices are dispensable aids to theory generation and, at worst, a distraction to logical ordered thought (Duhem, 1954). In a similar vein, analogical reasoning has commonly been presented as a variant of enumerative induction (e.g., Mill, 1872) that is incapable of supporting logically valid inferences and, therefore, is properly restricted to the context of discovery. Such views clearly indicate a marginalized role for models and analogical thinking in scientific inquiry. That is, by stressing the temporary nature of a model's influence on the inquiry process and by consigning analogical reasoning to the realm of psychological discovery techniques, these views

² A notable exception is Campbell (1920).

endorse the assumption that modelling practices are extraneous to the business of science proper.

Subsequent detailed analyses of scientific models and analogy, however, have overturned this assumption. In its place, philosophers such as Mary Hesse (1966) and Rom Harré (1976, 1978) have argued persuasively for the validity of analogical argument and for the centrality of modelling to scientific inquiry. More recently, this picture has been complemented and extended by a significant body of work in cognitive science (e.g., Forbus, Gentner & Law, 1994; Holyoak & Thagard, 1995; Magnani, Nersessian & Thagard, 1999). The richer representations of models and analogical reasoning afforded by these current cognitive accounts reinforce the importance that Hesse and Harré place on modelling practices for science. Moreover, by attempting to codify the cognitive processes involved in constructing analogies, investigating the ways in which such analogies advance our knowledge of poorly understood scientific phenomena, and developing ideas about how we can evaluate the effectiveness of our analogies, this research goes some way to uncovering the features that mark a productive scientific analogy. A brief summary of the basic principles of analogical thinking highlighted by these accounts can be given as follows.

Basic principles

Both philosophical and cognitive accounts of analogical reasoning begin by making a fundamental distinction between two components of an analogy: the *source* and its subject or *target* (Dunbar, 1999; Harré, 1976; Holyoak & Thagard, 1995). The source is an entity, process, or system, sometimes drawn from another domain, which is already well known and understood. The target is the unfamiliar phenomenon that the scientist is trying to understand. The importance of clearly distinguishing these two components of an analogy lies in the structured relationship that is established between them during the analogical reasoning process.³

³ For example, Harré (1976) has argued that this distinction is necessary for the role models play in creative theory construction.

The process of reasoning by analogy involves establishing a correspondence or *relation* between the source and the target that enables the transfer of information from one to the other. In practice, this is regularly achieved by constructing a model that draws on certain features of a known source to develop a plausible analogue of the primitively understood target. By mapping information about the nature of the source onto the target, that is, by viewing the target in terms of the source (Holyoak & Thagard, 1995), an attempt is made to learn something new about its nature. Viewed in this way, reasoning by analogy is essentially a method for developing knowledge about unfamiliar or novel phenomena by extension from what is already known (Hesse, 1966; Nersessian, 1999; Holyoak & Thagard, 1995).

Concerning the issue of what is mapped from source to target, accounts of analogy argue for a *selective emphasis* on particular features or properties rather than the wholesale mapping of a source in its entirety. For example, Hesse (1966) proposes that in any particular case the source is unlikely to correspond to the target in all respects. She therefore makes a distinction between what she terms “the negative analogy” (properties that belong to the source but not the target), “the positive analogy” (properties that are shared by source and target), and “the neutral analogy” (properties which need to be investigated to determine whether they form part of the positive or negative analogy). According to Hesse, the process of model building involves discarding the negative analogy and focusing on the positive and neutral analogies; the former providing an initial basis for comparison, and the latter offering the potential for new predictions (Hesse, 1966).

Similarly, cognitive accounts of analogy (e.g., Holyoak & Thagard, 1995) stress the selective mapping of properties that occurs in analogical reasoning. They further highlight the processes underlying this ability to extract features that are common to two or more situations while disregarding features that are different, as a significant cognitive advance that enables the transfer of information between two entities in the absence of global similarity (Holyoak & Thagard, 1995).

The capacity for selective mapping also underscores the ability to construct mappings between the source and the target at a number of different levels. Accounts of analogy typically distinguish between *attribute mappings* that focus on the basic

attributes or features of the source and target, and *relational mappings* that focus on the underlying patterns of relations between elements (Gentner, 1983, 1989; Hesse, 1966; Holyoak & Thagard, 1995). The most satisfying analogies are those that go beyond superficial perceptual similarities to uncover deep correspondences between the source and the target. For example, both Hesse (1966) and Harré (1976) stress the importance of mapping causal relations when reasoning by analogy in science. Similarly, Nersessian's (1999) discussion of analogical modelling focuses on the process of generic abstraction seen in classical mechanics, in which there is a move from specific examples of motion with all their attendant features, to the construction of models at increasing levels of abstraction that recognize common systems of underlying relations.

In addition to highlighting the components of analogy, the process of establishing correspondences between source and target, the issue of global versus selective mapping, and the levels at which mappings can occur, philosophers and cognitive scientists have been concerned to specify the *constraints* operating on analogical thinking. At a general level, researchers have argued that effective model construction in science is simultaneously constrained by both the source and the target. For example, Harré (1976) proposes that the model needs to maintain a relation with the source on which it is based. Yet at the same time it must also be adequate to the demands placed on it by the target, namely to provide a satisfactory explanation of the real object, process, or system under investigation.

More detailed analyses of the cognitive constraints that guide analogy use have identified *direct similarity* between source and target as a powerful constraint on establishing an initial correspondence between them and suggesting the possibility of more fundamental parallels (Holyoak & Thagard, 1995). The importance of *structural parallels* such as a one-to-one relation between elements of the source and elements of the target during the mapping process has also been emphasized (Holyoak & Thagard, 1995).

Finally, accounts of the use of analogy in science suggest that the fundamental purpose of analogical reasoning and model construction is to enable scientists to *move beyond perceptual experience*. For example, both Hesse and Harré argue that

modelling provides a important additional source of information to observation. Specifically, Harré claims that modelling is crucial to scientific endeavour precisely because it provides a means of reasoning about the existence and behaviour of a target's underlying causal mechanisms that cannot be investigated by other more direct methods (Harré, 1976, 1978).

At a more general level, both philosophical and cognitive analyses emphasize the ability of analogical thinking to effect *creative changes* to existing knowledge, and therefore stress the important role played by analogy in conceptual and theoretical innovation (Hesse, 1966; Harré, 1978; Holyoak & Thagard, 1995; Nersessian, 1999). However, at the same time these accounts recognize that analogy is a 'weak' method that does not guarantee a solution in any particular case. Accordingly, recommendations for the effective use of analogy in scientific inquiry stipulate that the output of analogical reasoning be rigorously evaluated against the real patterns and processes of nature (Harré, 1976).

The multiconstraint theory of analogy

The above sketch highlights some basic principles of analogical reasoning and provides an initial sense of how analogies can be employed to extend knowledge of unfamiliar phenomena. A far more detailed account that builds on these basic principles is the multiconstraint theory of analogy developed by Keith Holyoak and Paul Thagard (Holyoak & Thagard, 1995). This general cognitive theory, together with the associated computer programs it has inspired, provide an integrated treatment of analogical thinking that explains how people use analogy in everyday and scientific contexts in terms of the operation of three classes of constraints.

The first constraint highlighted by Holyoak and Thagard (1995) is *similarity*. When constructing an analogy between a source and a target, they argue that we are regularly guided by direct perceptual and/or semantic similarities between properties of the source and the target. For example, we may notice that two objects share similar features or that components of the source and target have similar functions or roles. According to Holyoak and Thagard, this natural 'overlap' between the source

and the target justifies establishing an initial mapping between them and provides a basis for investigating the possibility of more fundamental correspondences.⁴

The second class of constraints on analogical thinking concerns the *structural consistency* of mappings between the source and the target. According to the multiconstraint theory, the pressure to establish consistent structural parallels means firstly, that each element in the source should map onto a unique element in the target (there is a one-to-one mapping between elements), and secondly, that when groups of elements with particular relations holding between them are mapped from source to target, the relations holding these elements together should be preserved. When both these criteria are met, the analogy is said to constitute an isomorphism (Holyoak & Thagard, 1995).

Finally, Holyoak and Thagard go beyond earlier accounts of analogy that focus only on 'internal' issues of similarity and structure, to argue that analogy use is also highly constrained by the 'external' goals of the person using it. These goals provide the *purpose* of the analogy. Holyoak and Thagard (1995) highlight a range of purposes for analogy in scientific and everyday contexts including explanation, problem solving, decision-making, communication, and educational instruction. In each case, the purpose for which the analogy was originally constructed is shown to have a powerful impact on the subsequent reasoning process. Purpose, then, constitutes the third class of constraints on analogical thinking.

Importantly, the multiconstraint theory interprets the constraints of similarity, structure, and purpose as soft rather than inviolable constraints on analogy use that do not operate independently of one another. Rather, Holyoak and Thagard liken the function of these constraints to a complex of interacting pressures, some in agreement and some in opposition, the constant interplay of which pushes the reasoning process towards a satisfactory compromise. Successful attempts by Holyoak and Thagard to model this process as a constraint satisfaction problem in artificial neural networks,

⁴ As an illustration from the child-as-scientist literature, researchers began by noticing that children's conceptual structures demonstrated some of the features of framework theories in science, which led to a search for further connections between them (see for example Carey, 1985; Wellman, 1990; Wellman and Gelman, 1992).

have provided concrete demonstrations of how the multiple interacting constraints of similarity, structure, and purpose can work together to promote effective analogical thinking (Holyoak & Thagard, 1989b; Thagard, Holyoak, Nelson & Gochfeld, 1990).

Having specified the kinds of constraints that guide analogy use, and their interactive nature, Holyoak and Thagard (1995) develop an informative framework for representing analogies that identifies three successive levels at which correspondences between a source and a target can be established. These levels are differentiated from one another on the basis of increasing abstractness and complexity. The first and most concrete level concerns *attribute mappings*, that is, mappings between the basic attributes of the source and the target. For example, in comparisons of the development of knowledge in science and in childhood, “scientists” in the source domain can be taken to correspond to “children” in the target domain. Similarly, “scientists’ theories” can be seen to correspond to “children’s mental models”.

The second more complex level of correspondences identified by the multiconstraint theory involves *relational mappings*. In contrast to the basic attribute mappings highlighted above, relational mappings are concerned with similar relations holding between attributes in the source and target domains rather than the particular attributes themselves. Claims that children actively “construct” their mental models in the same way that scientists “construct” theories, is an example of a relational mapping: the relation between scientists and their theories, namely that of construction, corresponds to the proposed relation holding between children and their mental models.

A helpful way to express such propositions used by Holyoak and Thagard (1995) is

construct (scientists, theories)

where the predicate “construct” relates “scientists” and “theories” to one another by imposing the particular structure

construct (<constructor>, <constructed>).

Using this notation, we can express the relational mappings identified above as

construct (scientists, theories)
 construct (children, mental models)

which proposes that “scientists” are related to “theories” in the same way that “children” are related to “mental models”; in both cases the former “construct” the latter (Holyoak & Thagard, 1995, p.27; see also Shelley, 1999a, 1999b for informative applications of the multiconstraint theory to multiple analogies in archaeology and in evolutionary biology).

Finally, in addition to attribute and relational mappings, the multiconstraint theory identifies a third level of mappings that is both more complex and more abstract than either of the previous two levels. These mappings, referred to as *system mappings*, focus on “. . . the systemic properties of the source and target” (Shelley, 1999a, p.583), and involve mapping interconnected systems of higher order relations.⁵ Proposed system mappings suggested by advocates of the child-as-scientist analogy include the reasons or motivations underlying scientists’ theory construction and by analogy, children’s model construction. For example, Karmiloff-Smith (1988) has argued that both scientists and children construct theories in order to explain phenomena, which can be expressed as

in-order-that (construct, explain).

Alternatively, Gopnik and Meltzoff (1997) propose that evidence is the primary cause of theory change in science and, by extension, conceptual change in childhood; that is, evidence causes knowledge to change in both domains which can be expressed as

cause (evidence, knowledge change).

⁵ There are similarities between Holyoak and Thagard’s (1995) discussion of system mappings that centrally involve higher order relations such as “cause”, and Hesse’s earlier presentation of analogy in terms of (vertical) causal relations within analogues and (horizontal) similarity relations that map causal relations between analogues (Hesse, 1966).

Both these mappings provide examples of system mappings that can be differentiated from less complex attribute and relational mappings by the presence of higher order relations such as “cause”, and by the fact that it is *relations between relations* that are being mapped from source to target rather than basic attributes or first-order relations between attributes. Examples of all three levels of mappings discussed are presented in Table I.1.

Table I.1 The child-as-scientist analogy: Examples of attribute, relational, and system mappings.

	SOURCE	TARGET
Attribute	scientists scientific theories	children children's mental models
Relational	construct (scientists, theories)	construct (children, mental models)
System	in-order-that (construct, explain) cause (evidence, knowledge change)	in-order-that (construct, explain) cause (evidence, knowledge change)

A template

From the brief presentation given above, it is clear that the multiconstraint theory offers a comprehensive account of analogy and its role in creative thought. Firstly, by highlighting the operation of three classes of constraints on analogical thinking, the theory shows how it is possible to establish systematic correspondences between a source and a target. Secondly, by differentiating three levels of complexity in analogical reasoning, the multiconstraint theory demonstrates the capacity to represent even very complex analogies in a rich and meaningful way. As seen in Table I.1 above, this form of representation that parses an analogy into attribute, relational, and system mappings emphasizes the depth of the proposed parallels between source and target, and in doing so offers a basis for comparing competing analogies.

Finally, Holyoak and Thagard (1995) explicitly tie these increases in mapping complexity to the three constraints they have identified as fundamental to analogical thinking. Drawing on a wide range of examples, they show how more effective use of analogy is associated with a greater satisfaction of the constraints of similarity, structure, and purpose. This in turn lends support to their claim that reasoning by analogy can be profitably defined in terms of these multiple interacting constraints.

The above considerations suggest that the multiconstraint theory offers a useful template for representing the various child-scientist mappings that have been proposed by researchers, as well as providing criteria for evaluating the goodness or coherence of various formulations of the analogy. Using Holyoak and Thagard's (1995) theory, we can determine that a productive analogy between children and scientists will be one that successfully establishes system mappings involving higher order relations between the source and the target and does so by applying the constraints of similarity, structure, and purpose. In Chapter 7, I will employ the specific taxonomy of attribute, relational, and system mappings to undertake a comparative evaluation of the theory theory (Gopnik & Meltzoff, 1997), and my alternative methods-centred model of child-scientist parallels that I will develop through the course of this thesis. However, first it is necessary to examine Gopnik and Meltzoff's theory theory in more depth and to consider some of the criticisms that have been levelled against it. These tasks form the basis of Chapter 1.

Chapter 1

The Child-as-Scientist Analogy: Current Research and Applications

The central idea of this theory is that the processes of cognitive development in children are similar to, indeed perhaps even identical with, the processes of cognitive development in scientists. (Gopnik & Meltzoff, 1997, p.3)

In the Introduction, I highlighted three major claims by Gopnik and Meltzoff (1997): firstly, that children's conceptual structures are best described as theories; secondly, that their conceptual development is an ongoing process of theory construction and revision; and thirdly, that semantic development is effectively bootstrapped to this theory change process. Having provided a brief historical background to the theory theory, I now want to examine these claims in more detail by elaborating on the general framework Gopnik and Meltzoff develop in support of their proposals.

Accordingly, in Section 1.1 I provide a more in-depth account of the theory theory as presented in two recent formulations (Gopnik, 1996b; Gopnik & Meltzoff, 1997). Section 1.2 highlights a number of questions raised by this particular characterization of child-scientist parallels. In Section 1.3, I suggest that these problems demand a reorientation of the debate and outline my arguments for an alternative methods-centred model of cognitive change that I will develop in subsequent chapters.

1.1 The Theory Theory

It would seem helpful to begin discussion of the theory theory by returning to the key developmental questions posed at the beginning of this thesis. That is, in attempting to account for the development of knowledge in childhood, how does the theory theory characterize the content of children's knowledge, their reasoning processes, the impact of initial structures, the nature of conceptual change, and the role of experience in development?

Concerning the content of children's knowledge, the theory theory presupposes that children's conceptions of the world will be radically different from our everyday adult conceptions. In fact, it is a central tenet of child-scientist mappings that young children's knowledge of biological, physical, and psychological phenomena is embedded in distinctive frameworks that support different concepts and sets of relations amongst concepts from those found in adult theories (Carey, 1985, 1988; Wellman, 1990; Wellman & Gelman, 1992). This 'alternative models' view is strongly endorsed by Gopnik and Meltzoff (1997), who argue for a knowledge continuum on which children's theories are as far removed from adult everyday theories in their content as our everyday theories are in turn from current scientific knowledge.

While the theory theory predicts that children do know different things from adults about the world, it does not endorse the additional claim that this is due to immature reasoning abilities. Unlike Piaget's stage model, which essentially profiles a series of developmentally different thinkers (Wellman, 1990), a theory theory perspective 'uncouples' the differences found in children's theories from the development of their central thought processes. Gopnik and Meltzoff (1997) draw heavily on the analogy with science to argue that successive knowledge systems can exhibit qualitative conceptual change without corresponding changes in the underlying capacity for thought. From this standpoint, they reject the demarcation of child and adult cognition, and propose an essential continuity in the processes by which children and scientists develop knowledge of the world.

Regarding the question of initial structures, the theory theory in its current form advocates a strong nativist position whereby humans are endowed with a set of initial theories and a capacity for reworking these theories based on experience (Gopnik & Meltzoff, 1997). Gopnik and Meltzoff argue that this position is in general agreement with the wealth of data from recent work on infancy and early childhood, indicating that we begin life with a far richer set of representations than previously believed (e.g., Spelke, 1991). However, the theory theory diverges from standard characterizations of these structures as innate constraints, biases, or modules, by allowing them the same form, and importantly the same capacity for revision, as later knowledge (Gopnik & Meltzoff, 1997).

As indicated above, the theory theory is distinctive as an account of development because it characterizes innate knowledge as theoretical and therefore open to fundamental revision. Further contrasts with competing accounts become apparent when attention is turned to the nature of the developmental process itself. From a theory theory perspective, the pattern of change is inherently domain-specific as opposed to domain-general, and involves radical conceptual reorganization within theoretical frameworks rather than a gradual accretion of knowledge. The analogy with revolutionary theory change in science is fully exploited by Gopnik and Meltzoff (1997), who characterize development as “a succession of theories” and claim that theory revision often involves an entire theoretical framework. Moreover, semantic development is explicitly drawn into the ambit of the theory theory, with the emergence of specific words in a particular domain, and the identification of close relations between semantic and conceptual developments being tied to the underlying theory change process.

In this domain-specific restructuring account, children’s everyday experience acquires an evidential role. Gopnik and Meltzoff (1997) set forth a particular perspective on the interaction between children’s experience of the world and their developing representations that mirrors the relationship between empirical evidence and theoretical development in science. In their view, children’s everyday experience is properly classified as evidence for or against their developing theories, and as such constitutes a bona fide mechanism of developmental change. This classification indicates that the causal factors operating in scientific theory change are also

operating in cognitive development, and grounds Gopnik and Meltzoff's (1997) call for a unitary model of knowledge development to be applicable across science and childhood.

A consideration of the theory theory's stance on key developmental issues provides a useful overview of the main tenets of the theoretical framework. However, while versions of these ideas can be found in a number of earlier publications (e.g., Gopnik, 1984, 1988; Gopnik & Wellman, 1992, 1994), the latest defence aims to develop the theory theory beyond its current heuristic status. In a bid to achieve this goal, Gopnik and Meltzoff (1997) advance four main strategies.

1.1.1 Science and childhood: The cognitive connection

Their first strategy is to give science a cognitive characterization (see also Giere, 1988, 1992). In marked contrast to normative and sociological approaches, this characterization directs attention to the representational capacities and judgment strategies used by scientists to develop knowledge about the world, and therefore provides Gopnik and Meltzoff (1997) with a *common* platform from which to posit child-scientist parallels. Working from this platform, Gopnik and Meltzoff look to focus on the similarities between representations and rules underlying scientific knowledge and children's everyday knowledge and the processes that effect change in these representations and rules over time. Objections relating to phenomenology, the social structures of science, and the uniform development witnessed in childhood, are countered by arguing that none of these differences between children and scientists undermines this proposed cognitive connection.

Moreover, Gopnik and Meltzoff (1997) speculate that the close cognitive alliance between children and scientists may in fact have an evolutionary basis. According to Gopnik and Meltzoff, the key to uncovering this "common structure" lies in the basic theory formation capacities that have evolved in support of early childhood learning. They argue that these capacities not only play an important role in the growth of knowledge in childhood, but in a similar manner are also seen to underwrite scientific inquiry. In short, Gopnik and Meltzoff claim that the cognitive processes of

generating, developing, and revising theories that drive scientists' knowledge construction efforts have precursors in cognitive development. They further suggest that scientific progress can be made intelligible through these links with childhood learning. Working from a theory theory perspective, Gopnik and Meltzoff propose that the success of science is largely due to the evolutionary history of the cognitive mechanisms it utilizes. The reason science "gets it right" is not because of its peculiar organizational structures or social practices, but rather because it exploits cognitive abilities that have evolved to ensure early learning gets off the ground successfully.

1.1.2 Theories and theory change

The second strategy employed by Gopnik and Meltzoff (1997) is to offer a detailed account of scientific theories that identifies key structural, functional, and dynamic features. They argue that if the theory theory is to be developed, then claims that children hold theories and that cognitive development mirrors scientific theory change need to be far more specific. However, they also acknowledge that the absence of a definitive account of theories in the philosophy of science makes this task difficult. Their solution is to distil the following account of theories and theory change from a variety of philosophical sources (e.g., Kuhn, 1962; Hempel, 1965; Popper, 1965; Lakatos, 1970; Laudan, 1977).

With regard to structural features, Gopnik and Meltzoff (1997) argue that theories can be distinguished from other sorts of knowledge in four principal ways. Firstly, theories are abstract structures that postulate entities removed from, and underlying, the evidence they are generated to explain. Secondly, theories demonstrate a degree of interconnectedness that goes far beyond that of loosely grouped collections of facts or beliefs. Rather, the concepts and theoretical terms encompassed by a theory are more or less defined by their place in a web of constructs. Thirdly, theories identify, and make explicit, the causal relations holding between the theoretical entities and the observed patterns in the data. Finally, theories make specific ontological distinctions or commitments. That is, an essential feature of theories is they specify the kinds of entities that exist in the domain in question.

Concerning functional features, Gopnik and Meltzoff (1997) argue that theories are distinctive because of their predictive, interpretative and explanatory abilities. They allow predictions about a wide variety of evidence, not simply the evidential base that the theory was originally constructed to explain. Theories also provide a particular perspective on the domain in which they operate, including interpretations of the relevance of evidence. Finally, theories are unique because of their explanatory powers. They offer causal explanatory frameworks that enable their adherents to make sense of phenomena in the world.

The structural and functional features of theories listed above serve to differentiate theories from other knowledge structures. However, the most relevant features for Gopnik and Meltzoff's purposes are the dynamic processes involved in theory construction and revision. Working predominantly from a Kuhnian inspired account of scientific change, Gopnik and Meltzoff (1997) offer the following characterization of the epistemological processes involved in the shift from one theoretical framework to another.¹ Theory change is often preceded by the build-up of counter-evidence to the theory. To begin with, the importance of such counter examples may not be recognized and the evidence is ignored. Eventually, the need to account for the evidence is acknowledged and a collection of auxiliary hypotheses are developed for this purpose. However, under the weight of these additional hypotheses the theory gradually loses its earlier simplicity and coherence. A final phase in the transition requires the construction of an alternative view that is often selectively applied at first to those areas where the original theory proved insufficient. Only later is it recognized to provide a coherent account of both the anomalies and the evidence accommodated by the earlier theory. During this transition, a period of relatively

¹ In an earlier presentation of this account of theories and theory change (Gopnik & Wellman, 1994), Kuhn (1962) is cited (along with Lakatos (1970) and Laudan (1977)), whereas in Gopnik and Meltzoff (1997), despite the characterization being virtually identical to that given in Gopnik and Wellman (1994), the authors cite Kuhn (1977). However, given the essential continuity in the *content* of Gopnik and Meltzoff's account of theory change with the earlier formulation (see also Gopnik & Wellman, 1992; Gopnik, 1996), specifically its focus on the pattern of change described by Kuhn (1962) where a proliferation of anomalies are ignored, then accommodated by auxiliary hypotheses, until eventually a new theory emerges that is incommensurate with the earlier theory, the critique that follows focuses on the authors' reliance on Kuhn's account of revolutionary theory change as a source model for understanding cognitive development.

atheoretical experimentation and observation may occupy centre stage and provide an impetus for change.

1.1.3 Alternative accounts

A third strategy employed by Gopnik and Meltzoff (1997) to develop the theory theory involves contrasting it with alternative accounts of the structure and development of knowledge. They focus on two important contrasts: the first being with innate modules, and the second with what they collectively term “empirical generalizations”. In each case, they argue that the theory theory can be clearly differentiated from these alternatives along a number of dimensions.

Regarding innate modules, Gopnik and Meltzoff (1997) suggest that the key differences relate to the dynamic features of theories. From a modularity perspective (e.g., Fodor, 1983), development is seen to involve the maturation of innate modules, with experience typically allocated a minor ‘triggering’ role. The representations that result are in a strong sense predetermined and are not open to fundamental revision via evidence. Gopnik and Meltzoff argue that this picture of development contrasts markedly with that proposed by the theory theory, according to which children’s theories are inherently defeasible, and their experience functions as evidence in the knowledge development process.

While the contrast with modules is primarily focused on theory dynamics, the contrast with empirical generalizations looks to the structural and functional dimensions of theories. Gopnik and Meltzoff (1997) argue that theories can also be distinguished from a variety of knowledge structures that are loosely grouped under the heading “empirical generalizations” because of their common ties to immediate experience. Examples given are scripts, narratives, and connectionist nets.² According to Gopnik and Meltzoff (1997), the organization of these knowledge structures and the ways in which they function provide a significant contrast with theories. In particular,

² Gopnik and Wellman’s (1994) characterization of simulation accounts of the child’s developing understanding of mind would seem to fit into this category as well.

empirical generalizations do not show the abstractness and coherence typical of theoretical structures, and they offer only limited support for prediction, interpretation, and explanation.

1.1.4 Empirical defence

As a final strategy to develop their general explanatory framework, Gopnik and Meltzoff (1997) turn to a detailed empirical defence of the theory theory. Focusing on the period from infancy to early childhood, they chart the specific sequence of conceptual and semantic developments in three important domains: the understanding of appearances, actions, and kinds. In each case, they argue the theory theory makes the most sense of the developmental data by providing a coherent framework for explaining the concepts, words, inferences, and correlations between conceptual and semantic changes that characterize early cognitive development.

Specifically, their discussion of the evidence suggests that within each domain a continuous line of development can be traced through the non-verbal infancy data, recordings of children's spontaneous utterances, and later experimental work with preschoolers. They argue that this continuity is typically overlooked by conventional analyses that focus on the presence or absence of single concepts in children of a particular age. Moreover, Gopnik and Meltzoff propose that a theory theory analysis is able to deal successfully with seemingly incongruous evidence of both continuity and dramatic change in children's knowledge, by proposing an ongoing process of theory formation and revision that extends from infancy right through to adulthood.

Gopnik and Meltzoff's detailed examination of children's understanding of appearances, actions, and kinds, also allows them to consider evidence for domain-specific versus domain-general developments, and their discussion suggests that children's knowledge in these three areas may emerge relatively independently of one another. Finally, a great deal of Gopnik and Meltzoff's empirical defence involves mapping the relations between particular conceptual and semantic developments in each of the three domains, in terms of both their timing, and the rationale behind their joint appearance (i.e., why a particular word might plausibly be tied to a particular

conceptual development). They argue that the empirical relationships uncovered between specific words and concepts provides convincing evidence that theory changes lie at the heart of both cognitive and semantic development. That is, as in science, theory and language change in step with one another.

1.2 Questioning the extant framework

Gopnik and Meltzoff's (1997) presentation of the theory theory currently stands as the most detailed account of child-scientist parallels in the cognitive developmental literature. It also embodies a particularly strong version of the claim that continuities exist between scientific and lay cognition. Precisely because of this detailed and immoderate stance, the theory theory raises a number of questions that would not be prompted by weaker theoretical positions. In this section, I focus on three main areas of concern highlighted by critics: 1) the nature of the relations being established between science and childhood, 2) the extension of the theory theory to infant cognition, and 3) the source model of scientific change that underpins Gopnik and Meltzoff's formulation of the analogy.

1.2.1 Children and scientists: Analogy or identity?

An initial issue that relates directly to the strength of Gopnik and Meltzoff's (1997) claims is whether they are arguing for analogy or *identity* between children and scientists. Gopnik and Meltzoff do not clearly state their position on this issue and blur the distinctions somewhat by arguing at different points throughout their defence that both sorts of relations hold (Downes, 1999). However, a close look at their arguments for the theory theory indicates an endorsement of identity, or something extremely close to it, at the cognitive level:

Scientists and children both employ the same particularly powerful and flexible set of cognitive devices. These devices enable scientists and children to develop genuinely new knowledge of the world around them. (Gopnik, 1996b, p.486)

Our claim is that quite distinctive and special cognitive processes are responsible both for scientific progress and for particular kinds of development in children. (Gopnik & Meltzoff, 1997, p.49)

As indicated in Section 1.1, this claim for deep cognitive similarities between children and scientists is not seen to be jeopardized by differences in phenomenology, socialization, or timing. Moreover, an evolutionary argument is marshalled to explain why such cognitive similarities exist:

Our hypothesis . . . is that the most central parts of the scientific enterprise, the basic apparatus of explanation, prediction, causal attribution, theory formation and testing, and so forth, is not a relatively late cultural invention but is instead a basic part of our evolutionary endowment.

. . . we can think of organized science as taking natural mechanisms of conceptual change, designed to facilitate learning in childhood, and putting them to use in a culturally organized way. (Gopnik & Meltzoff, 1997, pp.20-21)

However, if this assessment of cognitive identity is an accurate representation of Gopnik and Meltzoff's position, then it raises a number of questions concerning the relations between both scientific and lay cognition, and child and adult thought. The first, highlighted in different ways by Giere (1996), Faucher, Mallon, Nazer, Nichols, Ruby, Stich and Weinberg (2002), and Stich and Nichols (1998), is whether attributing common theory formation capacities to children and scientists is *sufficient* to account for scientific development. For example, Faucher et al. (2002) point out that Gopnik and Meltzoff take the same theory revision process – innate theory revision mechanisms operating on existing theoretical structures in the face of competing evidence – to be at work in childhood and in science. By focusing exclusively on these three elements, they effectively rule out all other candidates vying for a significant role in scientific change including social and political processes (Stich & Nichols, 1998), instrumentation, symbolic notation, printing, and experimental methods (Giere, 1996), and culturally transmitted norms (Faucher et al., 2002). Yet, as Faucher et al. and others have shown, this move yields a seriously

incomplete account of scientific inquiry. Moreover, it makes it difficult to see how the theory theory could illuminate scientific cognition beyond its illumination of human reasoning processes more generally. As Giere (1996) puts it, “. . . [the theory theory’s] implications for the study of science as a human activity . . . seem limited to its implications for normal adults” (Giere, 1996, p.541).

Secondly, if the claim is for identity between children and scientists, then where does this leave adult lay cognition? Gopnik and Meltzoff’s argument for deep cognitive relations between children and scientists that somehow ‘bypass’ our commonsense adult inquiry practices sits uncomfortably with the more general thesis of continuities between everyday cognition and scientific cognition on which the theory theory depends. Indeed, proposing a closer link between science and childhood cognition than between science and adult lay cognition implies a peculiar sense of *discontinuity*. This impression is reinforced by Gopnik and Meltzoff’s suggestions that adult everyday cognition is a ‘starting state theory’ for science in the same way that infant’s innate theories form the starting state for later theories in childhood, and that children and scientists, but not lay adults, are active theorizers. However, this picture of inaction and atrophy in lay adults not only clashes with the continuity thesis, it also makes assumptions about inter-theoretic relations and the state of our commonsense explanatory frameworks that seem at odds with the realities of folk theoretical practice.³

Thirdly, if the cognitive devices responsible for key developments in childhood are identical to those responsible for theory change in science, then this forces a radical reconstrual of the cognitive development process. In particular, Gopnik and Meltzoff suggest that there will be little to distinguish the learning procedures of young

³ To take folk psychology as one example, McCauley (1986) has argued that it is unlikely to be the case that neuroscientific explanations of human behaviour will eventually eliminate our belief-desire theoretical framework; rather, some form of co-evolution seems probable (see also Chi (1992) for empirical findings concerning intuitive physics that support this claim). Moreover, if lay adults are ‘inactive theorizers’, then this would presumably mean that our folk theories have remained essentially static and non-progressive. Again using folk psychology as the example, this is clearly not the case. We only need to consider our current views of psychopathology and compare them with earlier conceptions, in order to see that significant conceptual development and change has occurred over time (e.g., Kemp, 1990).

children and adult scientists, and this theorizing account of knowledge acquisition applies from the beginning of life.

1.2.2 Modularity, maturation, and baby theorists

In Section 1.1, I indicated that a major tenet of the theory theory is to reject any fundamental demarcation of earlier and later thought in favour of an essential continuity in the processes by which children and adults develop knowledge of the world. A clearly stated aim of Gopnik and Meltzoff's (1997) formulation is to extend this continuity thesis to cognitive development in infancy and early childhood, thereby going beyond existing applications that have looked to investigate theory construction abilities in older children:

If the theory is supposed to answer Socrates' question, that is, to account for our general capacity to develop new knowledge, it should apply more generally and be true from the beginnings of life. (Gopnik & Meltzoff, 1997, p.4)

With this in mind, Gopnik and Meltzoff (1997) argue that we are endowed with rich innate theories of object appearances, human actions, and object kinds. The evidence that they present is taken to show that these "starting state theories" are very different in content from the commonsense theories that we subscribe to as adults, and this 'distance' is argued to be an indicator of substantial conceptual change. Most significantly, these theories are seen to be open to revision in the face of competing evidence from birth. Therefore, in looking to apply the theory theory to infancy and early childhood, Gopnik and Meltzoff are not only arguing for "theories all the way down", but for theory *revision* as well:

On the starting-state view, the child is innately endowed with a particular set of representations of input and rules operating on those representations. According to this view, such initial structures, while innate, would be defeasible; any part of them could be, and indeed will be, altered by new evidence. We propose that there are innate theories that are later modified

and revised. The process of theory change and replacement begins at birth. (Gopnik & Meltzoff, 1997, p.51)

However, this proposal for theories and theorizing in infants can be seen to raise a number of issues concerning the origins of knowledge and the impact of developing cognitive resources on children's abilities to acquire knowledge. To begin with, Gopnik and Meltzoff (1997) explicitly contrast their "starting-state nativism" with "modularity nativism", the view that knowledge is the product of innate modules and cannot be revised or overturned in response to evidence. As indicated in Section 1.1, it is this indefeasibility of modules that Gopnik and Meltzoff argue provides a key contrast with innate theories, and gives license to their claim that theory revision occurs from birth. Yet, as Stich and Nichols (1998) point out, nowhere in their book-length defence of the theory theory do Gopnik and Meltzoff provide direct experimental evidence that the processes driving the development of these starting state theories are rational responses to evidence. As it stands, Stich and Nichols argue that later emerging conceptual structures could be the result of either an evidence driven process or "... a modular system with a sufficiently varied set of parameters and triggers" (Stich & Nichols, 1998, p.443).

Secondly, not only do Gopnik and Meltzoff (1997) fail to discount a modularity explanation for the early developments they consider, but their extension of the theory theory to newborns further serves to highlight what many critics already regard as a significant source of *dis-analogies* between the child and the scientist. As Nersessian (1996) remarks in her commentary on Gopnik (1996b), it is an open question how maturation complicates the child-as-scientist analogy, and this point would seem to gain further purchase when the theorizers in question are 42-minute-old babies as claimed by Gopnik and Meltzoff (1997).

Moreover, by arguing that even newborns are active and competent theorizers, Gopnik and Meltzoff disregard a significant body of research that details the impact of maturational-driven changes in information processing capacities on cognitive performance. To take just a few examples, studies investigating the role of inhibitory control (e.g., Russell, Jarrold & Potel, 1994), working memory capacity (e.g., Keenan, Olson & Marini, 1998), and the capacity to construct hierarchies (e.g., Perner, 1991),

have all found significant correlations between conceptual developments (e.g., the development of a theory of mind) and these basic information processing capacities, suggesting that changes in these cognitive resources contribute to children's capacity for theory construction.

For the knowledge developments they review, Gopnik and Meltzoff explicitly reject the proposal that non-conceptual development plays a significant role in cognitive change, and argue "... relevant evidence, rather than theory-independent maturational or information-processing changes is responsible for the changes in the children's understanding of the world" (Gopnik & Meltzoff, 1997, p.185). Yet by ignoring the importance of maturational issues as they apply to the developing mind/brain, and hence to the young child's ability to theorize, the theory theory neglects some fundamental differences between the situation of the infant and that of the adult scientist. In my view, even more serious for the theory theory than its lack of integration with research on domain-general developments is the absence of any real mechanisms of developmental change.

1.2.3 Kuhnian revolutions and cognitive mechanisms of change

In the Introduction to this thesis, I indicated that a central criticism of the theory theory approach to cognitive development has been that the characterization it affords actively *inhibits* any useful discussion of developmental mechanisms. As Russell (1992) and Gellatly (1997) have argued, the idea that theory change is a model of cognitive development does not provide any clear directives about the sources of change. Likewise, Nersessian (1996) remarks:

Precisely because the advocates of the "theory theory" are not Piagetian with respect to the processes of cognitive development and conceptual change, they owe an account of the nature of the processes through which theories form and change, i.e., of what the activity of "theorizing" comprises. (Nersessian, 1996, p.544)

The expressed aim of Gopnik and Meltzoff's latest defence of the theory theory is to rectify this problem by specifying the dynamics of theory change that they claim underpin knowledge development in science and childhood. In this section, I consider how successful Gopnik and Meltzoff (1997) have been in achieving this goal. In particular, I question the ability of their research programme *in its current form* to deliver a fruitful characterization of knowledge acquisition in childhood.

Gopnik and Meltzoff (1997) acknowledge that when comparing conceptual change in children and theory change in science, the typical strategy of developmental psychologists has been to 'buy in' classical descriptions from the philosophy of science and utilize them to interpret the empirical data. As indicated in Section 1.1, they continue to endorse this approach elaborated by Gopnik and Wellman (1992), in which the child's development of an understanding of mind is seen to follow the same sequence of progression as the pattern of scientific theory change identified by Kuhn (1962). In particular, children are perceived to disregard anomalies initially, then resort to an ad hoc auxiliary hypothesis to deal with the counter evidence, use a new theoretical idea in restricted contexts, and finally restructure their folk understanding so that a new theory emerges that is incommensurate with the earlier theory (Gopnik & Wellman, 1992). While Gopnik and Wellman focus on highlighting these characteristic features of scientific theory change in the child's emerging conception of mind between the ages of two and a half and five years, Gopnik and Meltzoff (1997) want to take the analogy even further and consider what *motivates* this developmental process in the young child.

Working largely from Kuhn's account of scientific development, they identify (1) the accumulation of counter-evidence, (2) theory-internal simplicity or coherence demands, (3) the availability of an alternative model, and (4) a period of relatively atheoretical experimentation and/or observation, as the critical factors in scientific theory change. Having identified these factors in scientific development, Gopnik and Meltzoff attempt to superimpose them onto the empirical data of children's conceptual development. In particular, for each of the three domains they consider, Gopnik and Meltzoff focus on the relation between theory change and the accumulation of evidence that weighs against the specific theories that young children hold, and argue that changes in evidence *cause* theory changes:

[Concerning the child's theory of kinds] . . . [t]he theory theory proposes that the motivation for these changes comes from the infants' observations of the behaviour of objects. It is the result of evidence. (Gopnik & Meltzoff, 1997, p.185)

Such arguments rest on the assumption that there are substantive parallels between historical transformations in science and psychological transformations in children (Levine, 2000). More specifically, Gopnik and Meltzoff's proposals indicate that theories in science and those held by young children change for the same reasons, and hence, that the dynamics of scientific change identified by Kuhn (1962) can provide a source of causal-explanatory mechanisms for informing the study of cognitive development.

A major obstacle for this assumption however, is that Kuhn's theory of scientific development is not concerned with the scientist-as-individual. In a comprehensive review of Kuhn's ideas, Paul Hoyningen-Huene (Hoyningen-Huene, 1993) argues that Kuhn's focus is not on the cognitive structures and methods of analysis utilized by individual scientists to advance knowledge. Rather, his views of incommensurability and paradigmatic change are most accurately interpreted as a structural account of the development of science more generally. According to this account, it is the scientific community and not the scientist as an individual that is the 'agent' of scientific activity, whether this activity takes place within a tradition of normal science or occurs during a period of revolutionary change:

The agent of a scientific revolution is, like that of a tradition of normal science, a scientific community. This central thesis of Kuhn's is important above all for two reasons. First, an inquiry into the factors swaying theory choice in scientific revolutions amounts to an inquiry into the reasons behind the community's decision. In addition, the question of whether a given episode in scientific development should properly be ascribed to revolution or to normal science can only be answered relative to particular communities. (Hoyningen-Huene, 1993, p. 200)

Therefore, by drawing on Kuhn's account of scientific development as a source model for the dynamics of theory change in children, Gopnik and Meltzoff (1997) are appealing to terms and criteria that were devised to characterize social processes of change in the history of science.

Not surprisingly, when Gopnik and Meltzoff (1997) attempt to adapt Kuhn's revolutionary theory change model for their purposes, a number of difficulties and confusions arise (Downes, 1999; Bishop & Downes, 2002).⁴ The most significant of these for my purposes, is that Kuhn's (1962) account of revolutionary scientific change offers no insights about the process of theorizing as a *cognitive* activity. Hence, when Gopnik and Meltzoff propose that evidence causes theory change, and that simplicity or coherence demands play a major role in this process, they have no way of using the analogy with science to develop these proposals so that they become claims about cognitive mechanisms. As a result, Gopnik and Meltzoff offer no ideas about how information from the world might interact with children's conceptual representations to effect knowledge development. Similarly, they fail to indicate how children might monitor their representations of the world in terms of simplicity or coherence criteria.

In sum, Gopnik and Meltzoff (1997) want to claim that children construct theories and that the cognitive mechanisms underlying this theorizing process are identical to those used by scientists. As they say, ". . . scientists *must* be using some cognitive abilities to produce new scientific theories . . . What else could they be using?" (1997, p.15) They further recognize the need to give prominence to individual scientists' cognition and look to cast both children and scientists as cognitive agents in their account of developmental change. Yet their arguments for the theory theory are ultimately wedded to a source model of scientific change whose mechanisms of development are not readily translatable into cognitive/psychological terms. As a result, Gopnik and Meltzoff are unable to persuasively argue that the cognitive

⁴ For example, in a recent critique of the theory theory, Downes (1999) identifies eight separate theses concerning the relations between science and children's cognitive development, that suggest Gopnik and Meltzoff's defence of the theory theory conflates scientists' cognitive development with the historical development of science as a body of knowledge.

processes of theory construction and revision are common to children and scientists, and fail to offer a convincing account of the mechanisms responsible for cognitive development.

1.3 Refocusing the debate: A methodological perspective

The above critique of the theory theory has raised a number of questions about the relations between science and childhood, the application of the theory theory to infant cognition, and the inappropriateness of the source model of scientific development that underpins Gopnik and Meltzoff's formulation of child-scientist parallels. A number of researchers have taken such problems as sufficient reason to dispense with the analogy. In this thesis however, I propose a different route. Rather than discounting the utility of child-scientist parallels altogether, I intend, in the chapters that follow, to propose a detailed reformulation of the child-as-scientist analogy. I will argue that in looking to the scientific analogy to inform our investigations of children's knowledge acquisition, we need to select an account of inquiry that illuminates the *methodological* processes by which scientists, rather than the community at large, develop knowledge. Having indicated that a Kuhnian account of revolutionary change is unable to satisfy this requirement suggests that an alternative approach to developing the analogy is required, one that gives serious attention to the cognitive character of science, and provides a framework for promoting useful discussion of the knowledge development process. I will argue that recent developments in methodology, specifically the emergence of a general abductive theory of scientific method meets these demands, and when adopted as a source model for the child-as-scientist analogy, this theory has the capacity to illuminate the process of knowledge acquisition in childhood.

Accordingly, in Chapter 2 I present a general argument for a methods-centred model of cognitive change. I then turn my attention to selecting a specific methodological source model, and examine two orthodox theories of scientific method for this purpose. In Chapter 3, I consider a third alternative – a comprehensive abductive theory of scientific method – and argue for its adoption as an appropriate source

model for the scientific analogy. With this source model in place, Chapter 4 aims to establish meaningful relations between the abductive account of scientific method and the knowledge construction efforts of young children by defining a narrower role for the child-as-scientist analogy within an interactionist account of development. Working within this framework, compelling parallels are identified between the theory building strategies of scientists highlighted by abductive method and the data-to-theory moves uncovered in microgenetic analyses of children's problem solving. Having identified an abductive pattern of reasoning in children's spontaneous theory building, Chapter 5 extends the analogy to the processes by which children evaluate the quality of everyday explanations, drawing on a computational model of theory evaluation for this purpose. In Chapter 6, the focus turns to speculations regarding the cognitive origins of science, in light of the theory theory's appeal to an evolutionary warrant for child-scientist parallels. Finally, Chapter 7 summarizes the main arguments and reinforces proposals for the utility of a methodological perspective on the child-as-scientist debate by undertaking a detailed comparative evaluation of the theory theory and the abductive-methods account along a number of dimensions.

Chapter 2

Choosing a Source Model

In the previous chapter, I challenged Gopnik and Meltzoff's (1997) formulation of the child-as-scientist analogy, and suggested that a Kuhnian model of scientific change is incapable of moving the theory theory beyond its current heuristic status. I pointed out that Kuhn's (1962) general description of the rise and fall of scientific theories says nothing about the methods scientists use to generate, develop, and evaluate their theoretical constructions. Yet, an account of this theorizing process is vital if science is to provide a useful source model for thinking about childhood cognitive development.

Having rejected the model of scientific change that underpins the theory theory, I therefore begin my reformulation of the analogy by examining options for a more appropriate source model. With this in mind, Section 2.1 provides a general rationale for a methodological model, based on the centrality of method in science. In Section 2.2, I suggest that Gopnik and Meltzoff's specific attempt to develop the analogy via the interplay between theories and evidence in science requires a methods-centred approach. Section 2.3 then reviews the two major orthodox theories of scientific method as candidate source models for the analogy, and argues that neither account can offer an adequate portrayal of scientific inquiry. I conclude that if the child-as-scientist analogy is to be successfully developed beyond its current heuristic status, then an alternative theory of scientific method will need to be recruited.

2.1 Science as method

Given that throughout the remainder of this thesis I will develop a specifically methods-oriented perspective on the child-as-scientist analogy, an initial task is to defend this methodological reformulation of the debate. That is, what reasons do we have for characterizing science primarily in terms of method? And how does a science-as-method view help us develop the analogy beyond its current heuristic status? In an attempt to answer these questions and thereby provide a rationale for my proposed reformulation, I first present arguments from evolutionary epistemology supporting the centrality of method in science. I then turn to the specific claims for the theory theory made by Gopnik and Meltzoff (1997) and suggest that they can only be effectively developed within a methods-centred framework.

2.1.1 Evolutionary epistemology: An overview

Epistemology is the study of the origins, nature, and growth of knowledge. A minimal characterization of evolutionary epistemology provided by Campbell (1974) is that this study takes account of, and is consistent with, our evolutionary circumstances. Simply stated, this means that when we develop ideas about human knowledge, they are in accord with our status as products of biological and social evolution given by contemporary science (Campbell, 1974). Defined in this way, an evolutionary approach to epistemological issues can be seen to form part of the wider domain of naturalistic epistemologies. Such epistemologies interpret humans and human knowledge as natural phenomena that are the legitimate subject of scientific investigation, and hence recognize the relevance of scientific evidence for resolving epistemological problems. Within this broad framework, however, evolutionary epistemologists argue for a specific relationship between biological evolution and the growth of human knowledge. The exact nature of this relationship has been subject to a variety of interpretations.

Bradie (1986), for example, identifies two distinct interpretations concerning the application of biological evolutionary insights to the *mechanisms* versus the *content* of cognition. The first view, labelled the “evolution of cognitive mechanisms

program”, argues simply that the biological substrates that support cognitive activity have evolved through a Darwinian evolutionary process. The second view, referred to as the “evolution of theories program”, extends the application of evolutionary insights to the content of cognition itself by suggesting that there is an analogical relationship between the processes of biological evolution and the processes of knowledge development (Bradie, 1986).

An alternative classification system provided by Hooker (1989) forgoes the bipartite distinction between material mechanisms and cognitive content proposed by Bradie (1986) in favour of a continuum of positions that vary according to the strength of the relationship proposed between biological and cognitive evolution. Hooker distinguishes six positions along this continuum that range from the weak claim that evolutionary insights can be applied only to the biological substrates of particular cognitive functions such as perception, through the partial application of a natural selection analogy to the content of human knowledge, to a strong claim for a unified evolutionary model in which cognitive evolution is understood as “. . . a literal extension of biological evolution” (Hooker, 1989, p.105). According to Hooker, this last claim for identity between biological and cognitive evolution removes an artificial dichotomy between the evolution of the brain and the evolution of ideas inherent in the weaker claims. Moreover, by removing this distinction and replacing it with a thoroughly naturalistic view of human knowledge as simply the “cutting edge” of a single dynamic process, Hooker clears the way for the development of an adequate epistemology based on the evolution of complex regulatory structures (Hooker, 1989).

Hooker makes a persuasive case for his strong view of biological-cognitive relationships and the broad outline of his position will be adopted in the discussion that follows. However, if we accept that evolutionary insights extend in even a minimal (i.e., analogical) fashion to the content of human knowledge, then the evolutionary picture of human development can be seen to have important implications for our ideas of knowledge and knowing. Firstly, as a number of authors have emphasized (Campbell, 1974; Hooker, 1987, 1989; Popper, 1973), due consideration of our evolved status undermines the assumption that our interactions with the world have always been conducted from a position of knowledge and privileged understanding. Instead, our ‘starting state’ is revealed as one of complete

ignorance, including ignorance of ourselves and our surroundings, and even ignorance of what it is to know (Hooker, 1987). By taking ignorance rather than knowledge as a starting point, subsequent evolution can be seen to be total in every respect. As Campbell (1974) puts it “[w]e once ‘saw’ as through the fumbings of a blind protozoan, and no revelation has been given to us since” (Campbell, 1974, p.414).

Secondly, by rejecting a position of privileged understanding that sets us apart from other species and replacing it with one of ‘evolution from ignorance’, we are forced to recognize a fundamental continuity between ourselves and other species that extends to our knowledge-making efforts. If, as suggested above, human knowing has not been the subject of any special creation but has evolved along with all other components of life, then we would expect to find basic continuities with the processes by which other species interact with, and learn about, their respective environments. Regarding the form of this continuity, there is a consensus among a number of evolutionary epistemologists (e.g., Campbell, 1974; Hooker, 1987; Popper, 1973) that the fundamental process for achieving knowledge gains is best conceived as one of trial and error. This process is seen to be the only means available to creatures evolving from a state of ignorance (Hooker, 1987). Furthermore, its range of application is argued to extend from the simplest organism’s interactions with its surroundings to human scientific endeavour (Campbell, 1974; Hooker, 1987; Popper, 1973).

I have suggested that an epistemology compatible with the evolutionary view of human development is one that adopts ignorance as its starting point and recognizes a fundamental continuity in intelligent activity across species. What are the consequences of these evolutionary considerations for our understanding of science? In what follows I argue that their impact lies in the challenge posed to traditional views of our knowledge situation, both in relation to other species and to our position in the world. This challenge in turn demands a radical reconceptualization of scientific endeavour in which questions of method come naturally to the fore.

2.1.2 Implications for understanding science

To begin, by highlighting the belief that human understanding has evolved along with all other aspects of life, an evolutionary perspective is fundamentally at odds with traditional empiricist views of science, which characterize the scientific enterprise as a hierarchy of knowledge structures built on indubitable foundations. As Hooker (1987) has argued, if we recognize that nothing is given in advance, then empiricism loses its claim to such a guaranteed basis for knowledge and with it the rationale for erecting science, including its methodology, on top of this foundation:

If one knows in advance the nature of true knowledge (e.g., that it is grounded in sensory givens, or in mystical access to Platonic heavens) then an epistemology may be erected in these terms (using this knowledge!) and method deduced as the most efficient means to maximise knowledge thus understood. But if one knows nothing in advance, not even what it is to know, then the only thing that can matter is a study of the methods of relieving ignorance (Hooker, 1987, p.141).

From this evolutionary standpoint, it makes no sense to define science in terms of either direct knowledge received via the senses or a priori valid truths. Rather, Hooker (1987) argues that given the totality of our cognitive evolution and the absence of any 'givens', the only rational strategy is to focus attention on the methods by which we develop knowledge of the world. What Hooker shows then, is that by taking our evolutionary position seriously we are led away from a view of science built on privileged knowledge and towards one that is centrally concerned with method.

Moreover, by revealing the absence of any knowledge foundations on which to base the scientific enterprise, an evolutionary perspective not only supports a science-as-method view, but also indicates that an appropriate view of scientific methodology will be one that recognizes fundamental continuities with our commonsense modes of inquiry. As indicated above, the basic process of trial and error is generally recognized by evolutionary epistemologists to have broad application across nature, including application to human scientific endeavour. However while endorsing this

insight, some commentators such as Popper (1973) continue to assume an isolationist view of scientific inquiry, in which science is sharply demarcated from commonsense, and where scientific method is described in abstract terms.

In contrast, Hooker (1987) argues that if we are *consistent* in our application of evolutionary naturalist insights to questions of knowledge and knowing, then there is no justification for retaining a logical distinction between science and commonsense modes of inquiry. In particular, Hooker points out that science is not distinguished from other forms of human knowing by any privileged foundations that in turn dictate unique methodological practices for scientific inquiry. Moreover, he proposes that a scientific methodology that is suited to our evolutionary situation needs to frame questions of scientific method specifically in terms of its method *users*. For Hooker (1987), this means reinstituting the human scientist at the centre of science and defining scientific methodology not in abstract terms but in terms of our evolved cognitive and social capabilities for learning.

The above considerations suggest that by taking our currently best available scientific knowledge about our evolutionary origins seriously, we are led to a view of science in which methodological considerations are central. Further, an appropriate conception of scientific methodology is revealed as one that rejects any logical demarcation between the methods of science and everyday inquiry, and thereby encourages the development of a unified account of cognitive learning. Turning to the question of child-scientist parallels, a science-as-method view indicates that if we want to consider the ways in which children are like scientists, then it is both plausible and appropriate to locate our search for similarities at the methodological level. Arguments for the continuity of basic inquiry processes across science and commonsense add further weight to a methodological formulation of the debate by discounting any a priori distinction between scientific methods and commonsense methods and more generally supporting the search for interesting parallels between the knowledge construction efforts of children and scientists.

In this section, I have provided a general rationale for adopting a methods-centred view of science. However, in order to support my proposed reformulation of the debate, I need to show not only that a science-as-method view is an appropriate

perspective to take on science, but also that construing science in these terms can promote the development of the child-as-scientist analogy beyond its current heuristic status. In order to do this, I return to Gopnik and Meltzoff's (1997) attempt to extend the theory theory, and in particular, their focus on theory-evidence relations in science as the key to developing an explanatory account of cognitive development.

2.2 Theory-evidence relations: The key to explaining cognitive development

In Chapter 1, I argued that the Kuhnian model of scientific change underpinning the theory theory was an inappropriate source model for the analogy because it failed to offer any insights about the mechanisms responsible for cognitive development. Building on this point, I now want to highlight the potential utility of an alternative method-oriented approach by examining in more detail the nature of Gopnik and Meltzoff's (1997) argument for identification of theory change in science and cognitive development in childhood. With this in mind, I turn to the philosophical questions Gopnik and Meltzoff (1997) use to structure their discussion of the theory theory, as a way of clarifying the intended purpose of the analogy for cognitive developmental research.

2.2.1 Revisiting Socrates' and Augustine's problems

Gopnik and Meltzoff (1997) open their extended argument for the theory theory by posing two general philosophical problems. The first problem, attributed to Socrates, involves explaining our capacity to develop new knowledge. That is, given the concrete nature of our sensory experience, how is it possible to develop complex abstract representations of the world? The second problem, which Gopnik and Meltzoff attribute to Augustine, concerns our ability in early childhood to learn our first words. In particular, recognizing the arbitrary nature of the connections between

specific words and the things in the world that they represent, how do we manage to make these connections so quickly and so successfully?¹

Having selected these problems as a way of introducing the theory theory, Gopnik and Meltzoff (1997) align contemporary investigations of children's cognitive and linguistic development with these broader philosophical concerns, and propose that the analogy with science offers a solution to the questions posed by Socrates and Augustine, as well as providing an explanatory framework for the developmental data. In particular, by redefining cognitive development as theory change Gopnik and Meltzoff (1997) argue that the theory theory can be seen to chart a viable middle course between the extremes of rationalism and empiricism, allowing researchers to discard various versions of the claims that either we are born knowing everything (rationalism), or we learn everything (empiricism). The analogy with science, they suggest, offers the key to explaining how knowledge develops, and this key lies in the relations between scientific theories and evidence:

Scientific theory change is, after all, one of the clearest examples we know of the derivation of genuinely new abstract and complex representations of the world from experience. The model of scientific change might begin to lead to answers to the developmental questions and, more broadly, might begin to answer Socrates' philosophical question. (Gopnik & Meltzoff, 1997, p. 3)

... the answer we will propose to Socrates' question is also an answer to Augustine's question, and an answer that can explain the empirical facts about early meanings that were discovered by developmental psycholinguists. Thinking of cognitive development as theory change also gives us a new and better way of thinking about semantic development. (Gopnik & Meltzoff, 1997, p.6)

¹ For the original presentations of these problems, see Plato's *Meno* and Saint Augustine's *Confessions* (1.8).

In Chapter 4, I question Gopnik and Meltzoff's claims to have successfully defined a viable 'middle ground' for cognitive developmental research. The point to be made here, however, is that by weaving together philosophical questions about how we come to understand the world with findings from developmental research, Gopnik and Meltzoff (1997) can be seen to extend the analogy beyond earlier applications that focused primarily on descriptions of conceptual change (e.g., Gopnik & Wellman, 1992). Specifically, by placing the theory theory in direct competition with the theoretical frameworks of rationalism and empiricism, and by suggesting that it demonstrates the ability to provide answers to the philosophical questions raised by Socrates and Augustine, Gopnik and Meltzoff indicate they are concerned with developing an *explanation* of knowledge acquisition in the broadest possible sense. However, while the authors explicitly endorse this shift in the purpose of the scientific analogy, they continue to rely on a source model that at most offers them only a general description of the knowledge development process.

2.2.2 Theories and evidence: How does knowledge develop?

In order to collapse the distinctions between the development of knowledge in childhood and science, Gopnik and Meltzoff recast children's conceptual structures as theories, their experience of the world as evidence for or against these theories, and, as highlighted above, propose that children's development of successive representations of the world can be seen as the outcome of an interplay between the two. While this 'coalescence' gives them access to the resources used to explain scientific development, it also shifts the burden of proof onto this evidence-driven process. This in turn demands that their statements about the nature of the relations between evidential input and theories are not left open to interpretation but are instead anchored to an adequate theory of scientific method that provides them with a detailed explanation of how inquiry proceeds.

However, although Gopnik and Meltzoff recognize the explanatory demands their knowledge-driven account of development places on theory-evidence relations, they continue to rely on the following non-methodological account of scientific development, drawn largely from Kuhn (1962):

1. Counter evidence accumulates against the theory in question
2. Initially this counter evidence is disregarded
3. Ad hoc hypotheses are subsequently invoked to account for the anomalies
4. Over time these auxiliary hypotheses undermine both the theory's simplicity and its coherence
5. An alternative model to the original theory comes to light (often an extension of an idea that is already implicit in a peripheral part of the theory)
6. Initially the model is applied only to evidence not explained by the original theory
7. Eventually the new idea is adopted as a coherent explanation of both the anomalies and the evidence explained by the earlier theory
8. A period of intense experimentation and/or observation spans this theory change process (Gopnik & Meltzoff, 1997, p. 39).

This is the most detailed account of theory-evidence relations to be found in Gopnik and Meltzoff's (1997) defence of the theory theory. Indeed this description of the interplay between theories and evidence in science forms the template used by the authors in later empirical chapters of their book to characterize domain-specific developments as instances of the theory theory in action. Yet, in light of Gopnik and Meltzoff's explanatory goals, this list of "epistemological processes" would seem to be an inadequate source model for the theory theory in at least two fundamental respects.

Firstly, the general description of scientific theory change it provides fails to specify the crucial relations holding between theories and evidence, beyond saying that theories somehow come into contact with evidence and this 'contact', whatever it amounts to, somehow causes them to change. Secondly, this description says nothing about how the interplay between theories and evidence determines the truth or 'goodness' of scientific theories. Yet, Gopnik and Meltzoff rely on the view that science progresses towards an increasingly accurate account of the world, in order to

argue that theory *change* is a good model of cognitive *development*.² Gopnik and Meltzoff (1997) acknowledge this ambiguity to some extent:

Though the relation between the evidence and the change in the theory is, of course, far from simple, the theory theory proposes that there is something about the world that causes the mind to change, and that this fact ultimately grounds the truth of theories. (Gopnik & Meltzoff, 1997, p.53)

If the above assessment is correct, then the theory theory's focus has shifted from describing the content of children's knowledge and how it changes over time, to developing an explanation of the process by which this knowledge is generated, without undertaking a corresponding shift to a source model of scientific development that meets these new explanatory demands. Recognizing this fundamental mismatch between Gopnik and Meltzoff's (1997) model of scientific change and the purpose of the analogy, indicates the need for an alternative source model – one which offers a detailed reconstruction of the knowledge development process in science and, following Hooker (1987), frames questions of scientific method in terms of its method users. Adopting such a methods-centred account for the analogy would not only render the relations between theories and evidence explicit, but as will be shown in later chapters, also facilitate the investigation of common inquiry processes across scientists' knowledge-seeking endeavours and children's everyday problem solving.

Up to this point, I have focused on providing a rationale for adopting a methodological perspective on the child-as-scientist debate. This has involved reconstructing arguments from evolutionary epistemology that indicate the centrality of method in science, and demonstrating how the purpose of current attempts to develop the child-as-scientist analogy is consistent with a methods-oriented approach. Given the potential utility of a methodological source model for investigating child-

² The ambiguity in Kuhn's writings over the question of whether science progresses through revolutions (see Chalmers, 1999), would also appear to be at odds with Gopnik and Meltzoff's intention to use a broadly Kuhnian source model to explain children's cognitive progress.

scientist parallels, I therefore turn in the next section to examine two specific theories of scientific method as candidate source models for the analogy.

2.3 Orthodox theories of scientific method

Given the importance of theory-evidence relations to the explanatory arsenal of the theory theory, we can structure our search for an adequate theory of scientific method by asking how philosophers of science have traditionally interpreted these relations. That is, how do the different accounts on offer construe the interplay between the mind and the world? What roles have variously been attributed to perceptual experience and abstract conceptual representations in the construction of scientific knowledge? How does one constrain the other? Does information gathered from the world provide a foundation for scientists' theorizing about it and, if so, how? Finally, what mechanism/s of scientific method operate to ensure as far as possible that scientists are capturing the causal structure of the phenomena being studied? That is, what aspects of scientific method are responsible for ensuring the rigour of the knowledge construction process in science? These questions seem both important and useful if we are to settle on an adequate theory of scientific method that holds the promise of informing our theorizing about the development of knowledge in childhood.

2.3.1 An inductive account

A common perception of science holds that scientific knowledge is unproblematically derived from the facts of experience, and it is this view that lies at the heart of an inductive account of scientific method, expressed in the following quote by A. B. Wolfe:

If we try to imagine how a mind of superhuman power and reach, but normal so far as the logical processes of its thought are concerned, . . . would use the scientific method, the process would be as follows: First, all facts would be observed and recorded, without selection or a priori guess as to their relative importance. Secondly, the observed and recorded facts would be analyzed, compared, and classified, without hypotheses or postulates other than those necessarily involved in the logic of thought. Third, from this analysis of the facts generalizations would be inductively drawn as to the relations, classificatory or causal, between them. Fourth, further research would be deductive as well as inductive, employing inferences from previously established generalizations. (Wolfe, cited in Hempel, 1966, p.11)

According to this conception of scientific inquiry, often referred to as ‘naïve inductivism’, scientists have at their disposal a procedure for mechanically generating true theories about the world from a secure base of observable facts.³ Scientific inquiry proceeds by gathering all the facts in a disinterested manner to preserve their objectivity and then applying the rules of inductive inference to arrive at generalizations or laws concerning these facts. Precisely because of their solid basis in observation and the mechanical process by which they are derived from this foundation, these laws or theories are assumed to capture regularities actually existing in the world (see Figure 2.1).

From the description given, the inductive method seemingly offers the promise of an algorithm for truth. Critics of the inductive method however have identified a number of issues that undermine this promise. These issues concern both the nature of the facts from which scientific knowledge is supposedly derived, and the derivation process itself.

³ More recent attempts to articulate a viable inductivist account of scientific inquiry have softened the requirement that scientific knowledge be proven true. Instead, inductive arguments are given a probabilistic interpretation; that is, they are seen at best to lead to probable truth. For the sake of simplicity, the commentary that follows concentrates on a critique of naïve inductivism. However, most of these criticisms can also be made against a more sophisticated inductive probability account.

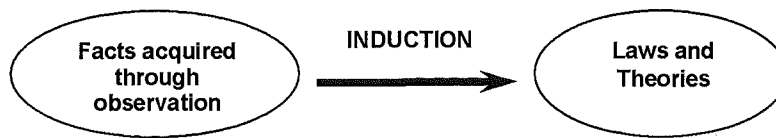


Figure 2.1 The main inferential move in an inductive account of scientific method.

Firstly, a number of commentators (e.g., Hempel, 1966; Curd, 1980) have argued that scientific inquiry on an inductivist view could never get underway if scientists were expected to collect all the facts while operating in a theoretical vacuum. Not only is an exhaustive recording of all facts impossible, it is also inappropriate. Scientists are typically interested only in those facts that pertain to the specific problem they are investigating. However as Curd (1980) makes clear, without a pre-existing theory to direct inquiry there is no way of determining which facts are relevant. Against the prescriptions of the inductive method, the observation and recording of facts will be necessarily circumscribed and guided by a scientist's theoretical preconceptions. This in turn means that these facts are to a certain extent dependent on the theories they presuppose (Chalmers, 1999). Acknowledging the theory dependence of the facts that constitute the foundation for science, then, suggests that this 'foundation' is not nearly as secure or infallible as the inductive method implies.

A second problem examined at some length by Chalmers (1999) concerns the issue of accessibility to these facts. The main reason facts are seen to offer scientific knowledge a secure base is that they are assumed to be directly established by the senses. This assumption, however, fails to recognize that a great deal of a scientist's time and energy is dedicated to active *intervention* in the world (Hacking, 1983). The rationale for conducting experiments would lose much of its force if facts relevant to science were readily observable and only needed recording. The reality of scientific practice where scientists regularly attempt to isolate the factors or processes under study and control for the effects of confounding variables speaks against such a

passive model. Rather it suggests that the facts of interest to science are typically abstracted from the world rather than 'given' in any straightforward sense.

The forgoing issues highlight problems with the inductivist view that directly observable facts provide a secure foundation for scientific knowledge. Even more problematic is the claim for a set of mechanical rules for inductively deriving true theories from this factual base. Not only does no such algorithm for truth exist, it is difficult to see how one could be conceived that would be capable of assuming the role it is allocated in the inductive account. To begin with, there are difficulties specifying what a satisfactory inductive argument actually amounts to (Chalmers, 1999). To illustrate, the principle of induction states:

If a large number of **A**'s have been observed under a wide variety of conditions, and if all those **A**'s without exception possess the property **B**, then all **A**'s have the property **B**. (Chalmers, 1999, p.47)

However as Chalmers (1999) points out, this meager characterization is not very helpful. It does not specify what "a large number" means, or indicate that this demand cannot be fixed in advance for all instances but will vary enormously depending on the particular problem under study. Similarly, this characterization provides no clues as to what constitutes a significant variation in conditions and on what grounds we are to make such a decision in any particular situation. Yet delimiting the potentially endless number of variations is essential if we are to reach a precise statement of the conditions under which a generalization constitutes a justifiable inductive inference (Chalmers, 1999).

A further point made by Curd (1980) indicates that important scientific discoveries involve "... an essential element of creativity and conceptual innovation which could never be performed by a machine following an algorithm" (Curd, 1980, p.207). The transition from data to a theory that delineates novel concepts not present in the evidential base seems to demand a certain degree of creative ingenuity. This suggests that even if a set of generally applicable and adequately specified rules of induction were made available, they would at most provide only a partial formula for generating scientific knowledge.

The problems with inductive method highlighted above, suggest the claim that science proceeds by mechanically deriving true theories from a secure base of observable facts, is untenable. The stipulation that scientists gather data in an undirected manner effectively means that inquiry proceeds in a blind fashion. The claim that scientific knowledge is inductively derived from these given facts is problematic because no one has been able to come up with an adequate account of this derivation process. Recognition of these difficulties has forced philosophers of science to consider alternative proposals for a workable account of scientific method. Not surprisingly, the successor to inductivism is an account in which the inductive generation of theories plays no part.

2.3.2 A hypothetico-deductive account

Carl Hempel, one of the primary advocates of the hypothetico-deductive account of scientific method, suggests that science proceeds by “. . . inventing hypotheses as tentative answers to a problem under study, and then subjecting these to empirical test” (Hempel, 1966, p.17). According to this view, also referred to as “the method of hypothesis”, a hypothesis or theory is indirectly tested by deriving from it a “test implication” or prediction via a deductive reasoning process. This prediction is then checked by observation or experiment. If the prediction is found to be false then the hypothesis is rejected. If the prediction is borne out by the test then, at least on Hempel’s view, it confers a degree of retrospective justification or confirmation of the original hypothesis (see Figure 2.2).⁴

In stark contrast to naïve inductivism then, the hypothetico-deductive model solely provides general operational guidelines for testing theories against the world. There are no prescriptions for generating theories in the first place. Neither Hempel nor Popper view discovery as a process of systematic inference that is amenable to logical analysis. Instead theories are seen to be “free inventions of the human mind” (Einstein, 1934), and therefore more relevant to psychology’s interests than to the

⁴ According to Karl Popper who argues for a falsificationist construal of hypothetico-deductive method, while successful tests of a theory’s predictions can be understood to offer some degree of corroboration for the theory, a theory can never be confirmed, only falsified (Popper, 1959).

philosophy of science. The essence of the inductivist programme is therefore dismissed as falling outside the bounds of scientific method. There is no logic to the discovery of theories, only to their subsequent confirmation.

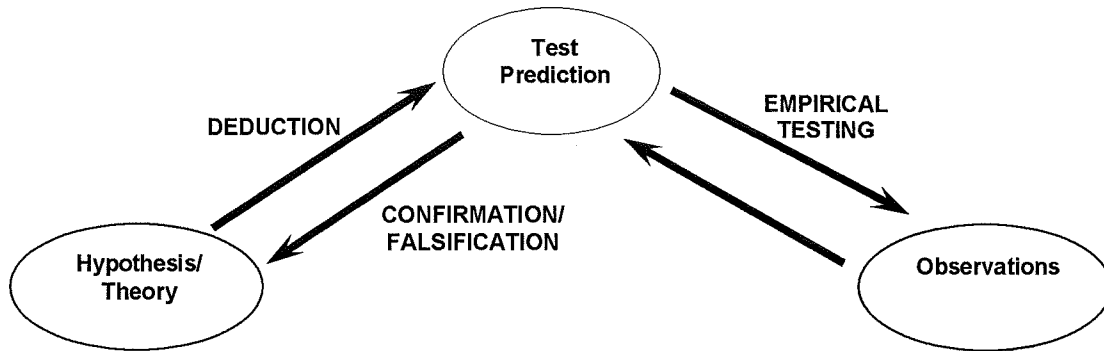


Figure 2.2 Hypothetico-deductive method.

At first glance, hypothetico-deductivism appears to provide scientists with a far more manageable formula for knowledge production. Appealing to deductive logic to derive predictions from a theory, and then checking these predictions for their empirical adequacy, invokes a (logically) rigorous test of a knowledge claim in terms of its consequences. However, despite claims to have achieved a workable account of inquiry by restricting science to a logic-and-testing exercise, this account of scientific method exhibits a number of serious failings.

Firstly and most seriously, is the inability of the hypothetico-deductive model to do what it claims to do, namely to adequately test the theories it seeks to evaluate. As a number of critics have pointed out (e.g., Glymour, 1980; Rozeboom, 1972; Salmon, 1967), in any realistic test situation as opposed to the pared-down examples used to exemplify a deductively valid argument, there will be a number of auxiliary assumptions or theories operating in conjunction with the theory under evaluation. This means that any predictions are not derived solely from the theory put forward for test, but from the *conjunction* of this theory and its associated auxiliary assumptions. In the case of an unfavourable test outcome, it will not be possible to identify whether

the responsibility for the failed test lies with the theory in question or with one of the auxiliary assumptions. Therefore, despite its promises to the contrary, the hypothetico-deductive account is unable to provide conclusive grounds for rejecting a knowledge claim.

Further limitations of this view of scientific method are revealed when we consider how little of the process of knowledge development in science it actually covers. In fact by confining methodological attention to theory testing, this approach leaves much of scientific inquiry unaccounted for. Most noticeably, hypothetico-deductivism disregards the methodology of theory generation, assuming that the initial “creative leap” to the theory is an activity that defies rational characterization. Accordingly, this account begins with the theory already formulated. However, such a disregard for discovery processes is both unhelpful and misleading. Specifically, it fails to acknowledge the existence of patterns of reasoning underpinning theory generation that indicate the strong possibility of a logic to discovery (e.g., Peirce, 1931-58; Thagard, 1988; Josephson & Josephson, 1994).

While the hypothetico-deductive model places theory generation outside the bounds of science, the methodological process of theory development is overlooked completely (Haig, 1987). The fact that this account begins with theory *testing* implies that theories are generated in a mature form. This assumption, however, fails to recognize the inherently developmental nature of scientific theories. As a result, this model of inquiry encourages the premature testing of undeveloped, often singular knowledge claims in science. Gains in the way of valuable knowledge from such a practice are, not surprisingly, negligible (Rozeboom, 1972).

Finally, hypothetico-deductivism presents an unduly restrictive notion of theory evaluation. The baseline assumption of this account is that testing theories for their empirical accuracy provides scientists with a conclusive test of a theory’s value. I argued above that in any actual scientific situation this assumption is unwarranted because the hypothetico-deductive model fails to test a theory in isolation from other knowledge claims. In addition, the recognition often referred to as the Duhem-Quine thesis, namely that theories will always be underdetermined by the available evidence (Duhem, 1954; Quine, 1963), indicates that any evaluation process based solely on

this criterion will be incomplete. Alongside predictive success, a range of super-empirical criteria including explanatory power, simplicity, fertility, and practical utility will need to be deployed. By demonstrating an exclusive concern with empirical adequacy however, the hypothetico-deductive account ignores this need to develop a multi-criterial perspective on theory evaluation.

In summary, while an inductivist conception of scientific method was shown to be untenable, a consideration of hypothetico-deductivism suggests that it fails to do justice to the actual structure of scientific reasoning. Relegating theory generation to the status of guesswork that involves no codifiable pattern of reasoning, ignoring theory development, and restricting the focus of evaluation to a theory's predictive success, all serve to reduce scientific method to a simple logic-and-testing model. By employing this reductive approach, a hypothetico-deductive account produces operational guidelines for what is in effect a highly circumscribed theory testing strategy. By itself, this strategy is insufficient to meet the demands of an adequate theory of scientific method.

2.4 Beyond orthodoxy

In this chapter, I have argued in favour of a methodological reformulation of the child-as-scientist debate, and have examined two orthodox theories of scientific method with a view to selecting an appropriate source model for the analogy. Specifically, I have evaluated each account in terms of the process by which it claims to give us knowledge of the world, and the various roles it assigns theories and evidence in this knowledge construction process.

According to a naïve inductivist account, scientists gather data and then derive true theories as output. This view stresses the primacy of observation; that is, information from the world provides an infallible and secure foundation for the theory generation process that follows. In contrast, a hypothetico-deductive account stresses the primacy of theory. According to its prescriptions, scientists begin with a theory already formulated and use deductive logic to derive predictions from this theory that

are amenable to empirical test. On both accounts then, the world constrains scientists' theorizing efforts, but in fundamentally different ways. While naïve inductivism advances a form of *foundational generation* where theories are constrained by the factual base from which they are derived, hypothetico-deductivism advocates *foundational justification*, in which the world is the reference point for evaluating scientists' guesswork and where theoretical constructions are only accepted into the body of scientific knowledge if they withstand the rigours of an empirical testing process.

A review of the range of criticisms that have been levelled at these two orthodox theories of scientific method by philosophers of science reveals that neither account is capable of providing an adequate perspective on scientific inquiry. This suggests that if we are to develop the child-as-scientist analogy along methodological lines, then we must adopt a more informative theory of scientific method as our source model. With this aim in mind, I move beyond orthodox accounts in the next chapter to examine recent developments in scientific methodology, and in particular the emergence of a comprehensive abductive theory of scientific method.

Chapter 3

An Abductive Theory of Scientific Method: A Third Alternative

In Chapter 2, I provided a general rationale for a methodological reformulation of the child-as-scientist analogy and argued that the investigation of child-scientist parallels would benefit from a methods-oriented approach. Specifically, I showed that Gopnik and Meltzoff's (1997) focus on theory-evidence relations in science needs to be tied to a codified account of the inquiry process. A review of the major two theories of scientific method, however, revealed that neither inductivism nor hypothetico-deductivism has the resources to provide an adequate portrayal of this process. In order to recruit a workable account of scientific method as a source model for the child-as-scientist analogy, it is therefore necessary to consider a third alternative. The alternative to be presented in this chapter, drawn from recent developments in methodology, is an abductive theory of scientific method (Haig, 2002).

In Section 3.1, I introduce the abductive theory by way of its dual focus on the detection and explanation of empirical phenomena, and highlight links to both a renewed philosophical interest in the empirical base of science and a growing recognition of abduction as an important species of scientific inference. Section 3.2 presents the abductive framework in more depth by detailing the multiple interacting contexts of inquiry that comprise abductive method. In Section 3.3, I conclude by briefly examining arguments for the domain and context specificity of methods in science and suggest that such arguments do not undermine the abductive theory's status as a general theory of scientific method.

3.1 Detecting and explaining empirical phenomena: Twin goals for inquiry

3.1.1 (Re)focusing on empirical phenomena

In the previous chapter, I identified a number of limitations facing inductivism and hypothetico-deductivism and suggested that neither account met the demands of an adequate theory of scientific method. An inductive or ‘bottom-up’ approach, as Chalmers (1999) terms it, places undue stress on the infallibility of observations and fails to demonstrate how scientific knowledge is derived from this secure foundation. In contrast, a hypothetico-deductive or ‘top-down’ approach side-steps the generation process altogether, locating the interplay between theory and evidence within a justificatory context. However, this received view of scientific method is unable to show how the process of checking a theory’s test predictions against the world affords a decisive assessment of a theory.

The failures of orthodox accounts of scientific method to adequately characterize the relationship between scientific theories and the material world, has led one group of philosophers to challenge widely held assumptions about the nature of empirical facts and their role in scientific inquiry. In their view, many of the problems confronting contemporary philosophy of science stem from an inaccurate rendering of the empirical base of science and can be resolved by redirecting philosophical attention to experimental practice. The proponents of this view are collectively known as the new experimentalists and the school of thought they have engendered as “the new experimentalism” (Ackermann, 1989).

The new experimentalism

According to Ian Hacking, one of the pioneers of the new experimentalism, contemporary philosophy of science is curiously anti-experimental in its outlook. This is odd, Hacking insists, because ‘experimental method’ used to be synonymous with ‘scientific method’ (Hacking, 1983, p.149). For example, during the scientific revolution of the seventeenth century, comments by the philosopher Francis Bacon

that scientists should ‘twist the lion’s tail’ – actively intervene in the world rather than merely observe it – were representative of, and in keeping with, scientific practice. In contemporary philosophy of science, however, such references to experimental manipulation have (until recently) been conspicuously absent (Hacking, 1983).

In an attempt to identify the historical reasons behind this shift to an anti-experimental stance in the philosophy of science, Hacking (1983) points to both the rise of positivism with its corresponding focus on observation, and the subsequent emergence of the theory dependence movement, as pivotal in the divorce of scientific method and experimentation. Concerning the impact of the positivist tradition, Hacking argues that prior to about 1800 (which he identifies as the starting point for positivism), the notion of observation was not central to philosophy of science. Yet after this date, observation, and in particular the creation of a fundamental *distinction* between observation and theory became increasingly important to discussions of science and scientific reasoning.

According to Hacking (1983), the creation of this distinction in philosophy of science altered what was understood to be real, with reality becoming confined to what could be observed with the unaided senses. Having this firm foundation of observational facts as a starting point, philosophical attention was redirected towards the development of increasingly sophisticated accounts of how theoretical statements could be logically derived from this observational base. This focus, however, meant that questions about the *production* of these empirical facts were virtually ignored. As Ackermann (1989) puts it, “[o]ne simply began to philosophize on the assumption that science was capable of delivering a data base of settled observational statements” (1989, p.185).

By ignoring the ways in which reliable observational facts are produced, positivism encouraged the view that the process by which scientists obtained such facts was straightforward and unremarkable. This view in turn offered little rationale for directing philosophical attention to issues of how scientists manipulated or intervened in the world, and consequently the perceived importance of experimentation to the task of characterizing scientific method was noticeably downgraded (Hacking, 1983).

While Hacking (1983) credits positivism with initiating a reduction in the perceived importance of experimental practice, he argues that it has been the subsequent theory-dependence movement that has been primarily responsible for its disappearance from philosophy of science discussions. Beginning with Kuhn (1962), this movement proceeded to dismantle the secure foundation of observable facts that had provided the starting point for positivist accounts of scientific knowledge. In its place, all observation was argued to be theory-laden to the extent that even the most commonplace observations were seen to depend to some degree on theoretical assumptions (e.g., Kuhn, 1962; Popper, 1959).

With theory dictating, or at least having an influential role in, observation the fundamental division between observations and theoretical inferences promoted by positivism was held to be no longer sustainable. However, in addition to sweeping away the positivist legacy of secure facts, Hacking (1983) argues that the currently dominant theory-dependence movement has also pushed experimental practice even further into the philosophical background. History of science, he suggests, has become the history of scientific theories, and the role of experimentation in this history is either downplayed, ignored or in some cases actively rewritten.¹ Moreover, it is accepted that theory precedes experiment both temporally and in terms of its overall importance to an account of scientific method. A quote from Popper attests to the relative weightings given to theory and experiment in contemporary philosophy of science:

¹ Hacking (1983) provides some interesting examples in which important experimental episodes in the history of science have been misrepresented as simply exercises in theory testing or confirmation, supporting his claim that experimental science is often rewritten as “theory history”.

The theoretician puts certain definite questions to the experimenter, and the latter, by his experiments, tries to elicit a decisive answer to these questions, and to no others. All other questions he tries hard to exclude . . . it is a mistake to suppose that the experimenter . . . [aims] ‘to lighten the task of the theoretician’, or . . . to furnish the theoretician with a basis for inductive generalizations. On the contrary the theoretician must long before have done his work, or at least what is the most important part of his work: he must have formulated his question as sharply as possible. Thus it is he who shows the experimenter the way. (Popper, 1959, p. 107)

This ‘theory-before-experiment’ view suggests not only that theory is logically prior to experiment in scientific inquiry, but that the only role of experiments is to test theories. In short, there is no *generative* role for experimental findings in scientific research. Rather, an experiment makes sense only if it is providing an answer to questions that are first put forward by theoreticians (Hacking, 1983).

By highlighting the rejection of pre-theoretical observations or experiments in theory-driven philosophy of science, Hacking (1983) shows how the theory dependence movement has shifted the onus of responsibility for scientific knowledge squarely on to the shoulders of theoretical conjecture. However, he points out that with this shift has come many of the problems that confront contemporary philosophy of science, most notably the absence of any secure base for scientific knowledge and the accompanying threat of disintegration into an anarchic state of ‘anything goes’ when the notion of theory-dependence is taken to its extreme conclusion (e.g., Feyerabend, 1975).

Against this backdrop, Hacking and other experimentally inclined philosophers argue for what Hacking terms a “Back-to-Bacon movement” in philosophy of science (Hacking, 1983, p.150; see also Franklin, 1986, 1990; Galison, 1987; Mayo, 1996). These new experimentalists reject outright Popper’s dictum that the only role of experiment is to test theories. In doing so, however, they do not attempt to resurrect the earlier positivist assumption that observational facts are unproblematically given by the senses. Instead, they argue that attending closely to actual experimental

practice offers an alternative to both traditions, in which a revised empirical base for science is located in the results of experiment.

From the detailed narratives of experimental episodes in the history of science constructed by the new experimentalists, three important themes emerge (Mayo, 1996). Firstly, the fine-grained analyses of what actually goes on in experimental practice indicate that understanding the role of experimentation in science offers a means of removing doubts about the objectivity of observation introduced by a theory-dominated picture of inquiry. In-depth discussions of the techniques and procedures employed in experimental research indicate the availability of a range of practical strategies and intervention tactics for reliably establishing experimental effects that do not depend on the application of high-level theory.

Secondly, the existence of such theory-independent warrants for determining the reality of empirical findings has led Ian Hacking to argue that “[e]xperimentation has a life of its own” (Hacking, 1983, p.150), and the results of such experiments constitute a body of controllable and reproducible empirical facts that do not disappear in the face of changing theory. Moreover, recognition of the stability and permanence of such experimental knowledge suggests an important form of (empirical) continuity and progress in science that is not apparent from a theory-dominated stance (Chalmers, 1999; Hacking, 1983; Mayo, 1996).

Finally, a consistent finding by the new experimentalists in their probing of experimental episodes is a focus of activity on what Deborah Mayo terms “the local discrimination of error” (Mayo, 1996, p.60). According to Mayo, these experimental narratives reveal that much experimental testing involves breaking a substantive problem or inquiry down in such a way that it is amenable to techniques that test for, and rule out, specific errors. Such a focus suggests it is the ability to differentiate real effects from artifacts, and thereby establish the existence of empirical phenomena, that is the cornerstone of experimental knowledge. What the new experimentalists have begun to develop with their detailed historical narratives of experimental episodes in science is a catalogue of methods and techniques for reliably achieving

such knowledge that has been overlooked by contemporary accounts of scientific inference focused on the appraisal of large-scale theories.²

To sum up, the impact of the new experimentalist movement in the philosophy of science has been to rekindle interest in the empirical base of science by providing a wealth of information about actual experimental practice, including the focus of researchers' efforts on obtaining reliable data, the experimental processes and techniques that are used to achieve this end, and the role that the resulting empirical knowledge plays in scientific inquiry. In keeping with these insights, a related body of work has attempted to refine ideas about the exact nature of this empirical base by introducing an important distinction between data and phenomena.

Data and phenomena

The popular or characteristic view of inquiry maintained by contemporary philosophy of science is that scientific theories explain and predict facts about what is observed. However, in a series of influential papers (Bogen & Woodward, 1988, 1992; Woodward, 1989, 2000), James Bogen and James Woodward argue that this assumption is fundamentally mistaken, and stems from a failure by philosophers of science to distinguish claims about *data* from claims about *phenomena*.

Bogen and Woodward describe phenomena as stable, general features of the world. Phenomena are the robust empirical regularities or 'effects' that scientists attempt to detect using a variety of experimental and statistical techniques, and explain by appealing to general theory. Bogen and Woodward point out that it is claims about phenomena that are the focus of scientific explanation and prediction and can serve as evidence for a theory under investigation. Examples of phenomena given by Bogen and Woodward include the melting point of lead, weak neutral currents, and recency effects in short-term memory (Bogen & Woodward, 1992).

² Mayo (1996) builds on the experimental narratives provided by Ian Hacking and others to construct a comprehensive philosophy of experiment that focuses on the process of validating claims experimentally by subjecting them to severe tests.

In contrast, data are public records or reports of discrete measurements or readings on a recording device (e.g., a thermometer, a questionnaire, etc.), that comprise the results of a particular experiment or study. Data are in a form that is accessible to the senses, and hence are open to public inspection and scrutiny. The publicly accessible nature of data is essential for the evidential role that data play in science. As Bogen and Woodward (1988) point out, data constitute the observational evidence for the existence of phenomena, which are typically unobservable. Data, therefore, provide researchers with their 'window' on phenomena. They (potentially) carry information about the existence of phenomena and as such are produced and interpreted by scientists for the sole purpose of extracting that information. Examples of data that might serve as evidence for the phenomena listed above, include reports of individual temperature readings in the case of the melting point of lead, bubble chamber photographs in the case of weak neutral currents, and reports of reaction times and error rates in certain psychological experiments in the case of recency effects in short-term memory (Bogen & Woodward, 1992).

Given the lack of attention to experimental practice in theory-dominated accounts of inquiry highlighted by the new experimentalists, it is perhaps not surprising that the data-phenomena distinction advocated by Bogen and Woodward has been ignored in philosophy of science until recently. However, in detailing the points on which phenomena and data differ, Bogen and Woodward (1988) discuss why it is a mistake to confuse or conflate the two and assume that science explains facts about observed data. Firstly, this assumption fails to acknowledge the proper objects of scientific explanation. While data are usually straightforwardly observable, they cannot be predicted or explained in any systematic way. This is because the data produced in an investigation will reflect not only the causal influences of the phenomenon under study (assuming it is successfully detected), but also a complex combination of causal factors that are peculiar to the specific experimental setup. Therefore, data will not possess the recurrent invariant features that make phenomena natural candidates for systematic explanation and prediction.

Secondly, the assumption that scientific theories explain and predict facts about what we observe obscures the evidential role that data play in science and with it the set of procedures scientists use to reason from claims about data to claims about phenomena – procedures that indicate scientists are primarily concerned with ensuring data *reliability*. Drawing on the narratives of important experimental episodes in science constructed by the new experimentalists, Bogen and Woodward (1988) highlight a wide range of these procedures routinely used in experimental practice including control of possible confounding factors, replications, statistical analysis, data reduction, and the empirical investigation of equipment including calibration of instruments. They argue that despite the popular view of science, none of these procedures are concerned with explaining the data. Rather, their aim is to identify and control for factors that could adversely affect the data's reliability. Therefore, sufficiently distinguishing data and phenomena not only accords with what scientists engaged in experimental work actually do, but also clarifies why they do it – ensuring the reliability of data is critical because it forms the grounds for claiming that phenomena exist (Woodward, 1989).

In summary, the refined understanding of what science should be attempting to explain and predict offered by Bogen and Woodward, together with the more general resurgence of interest in the empirical base of science promoted by the new experimentalists, reveals the neglect of experimental practice in philosophy of science and indicates the benefits of refocusing on the role of empirical phenomena as a solution to problems that have plagued theory-dominated accounts. For the purposes of this chapter, this work is instructive because it suggests that a more adequate theory of scientific method will need to incorporate an account of how reliable empirical facts are produced in science, and give greater attention to their role in scientific inquiry than orthodox accounts have previously allowed.

3.1.2 The role of abduction in scientific explanation

In the previous section, I focused on the detection of empirical phenomena and suggested the need to incorporate a role for such detection processes in a more satisfactory theory of scientific method than that offered by either naïve inductivism or the received hypothetico-deductive view. Giving due attention to how scientists manage to reliably *detect* phenomena and move from claims about data to claims about phenomena, suggests the need for a corresponding focus on how phenomena are *explained*, that is, an account of the forms of reasoning that allow scientists to move from descriptive claims about phenomena to explanatory theory. Drawing on insights from the new experimentalists together with the work of Bogen and Woodward, I indicated in the previous section that reasoning from data to phenomena is predominantly empirical in character. Here I draw on the work of the American philosopher Charles Sanders Peirce (1931-58) and those who have followed him, to suggest that reasoning from phenomena to theory is predominantly a matter of what has been termed *abductive inference*.

Introducing abduction

The term ‘abduction’³ was first used by Peirce to refer to the process by which explanatory hypotheses or theories are formed. According to Peirce (1934), abduction or “the operation of adopting an explanatory hypothesis” in order to account for some puzzling phenomenon, constitutes a distinctive kind of reasoning that is pervasive in science and in everyday life. To take an example, I was recently on a bus travelling to university, when I noticed an elderly woman who ran into the middle of the busy road ahead of the bus and frantically waved her arms up and down. With no conscious effort, I found myself thinking she must be trying to ‘flag down’ the bus. Abductively I formed the hypothesis that she wanted the bus driver to stop

³ Peirce also used the terms ‘hypothesis’ and ‘retroduction’ to refer to this form of explanatory reasoning. Other labels given to the process of forming an explanatory hypothesis (although differing from Peircean abduction in some respects), include ‘inference to the best explanation’ (Harman, 1965) and ‘explanatory induction’ (Rozeboom, 1972, 1990, 1997). In this thesis I use the term ‘abduction’ to reflect its increasing prominence in philosophy of science discussions (e.g., Niiniluoto, 1999) and widespread application in artificial intelligence (AI) research on problem solving (e.g., Josephson & Josephson, 1994), as well as its adoption in cognitive science (e.g., Magnani, Nersessian & Thagard, 1999; Magnani, 2001).

and pick her up, which provided a plausible explanation for her rather unusual (not to mention risky) behaviour.⁴

The above example demonstrates the ubiquitous human tendency to postulate motives or particular states of mind in order to explain the actions of other people. Further instances of everyday hypothesis formation are found in perception, for example when we are faced with incomplete or ambiguous visual stimuli that we need to make sense of. In science, researchers frequently construct hypotheses to attempt to explain puzzling facts such as observations by nineteenth century astronomers that the orbit of Uranus diverged from what was expected, or experimental findings by Lavoisier and others that objects gain rather than lose weight during combustion. In all these cases, abduction plays a pivotal inferential role (Thagard, 1988).

Peirce identified abduction as a valid form of inference that is importantly distinct from the classically conceived forms of induction and deduction. That is, he observed:

There is a large class of reasonings which are neither deductive nor inductive. I mean the inference of a cause from its effect or reasoning to a physical hypothesis. (Peirce, 1982, vol.1, p.180)

Moreover, Peirce's analysis of the inferences scientists make in reasoning to their hypotheses suggested to him that abduction exhibits a definite logical form:

The surprising fact, C, is observed;
But if A were true, C would be a matter of course,
Hence, there is reason to suspect that A is true. (Peirce, 1934, vol.5, p.117)

⁴ Presumably, the driver abducted a similar hypothesis to account for the woman's behaviour because he immediately pulled over, let her on the bus, and reproached her for risking her life in order to get him to stop.

According to Peirce, while deduction involves reasoning from a hypothesis to an empirically testable consequence that can be checked by experiment, and induction involves reasoning from a class of observations to descriptive generalizations based on those observations such as reasoning from a sample to a population, abduction is primarily an *explanatory* move in which we create a hypothesis or theory to account for observed phenomena by postulating the (potential) *causes* of those phenomena:

All the ideas of science come to it by the way of Abduction. Abduction consists in studying facts and devising a theory to explain them. (Peirce, 1934, vol.5, p.90)

The great difference between induction and hypothesis [i.e. abduction] is, that the former infers the existence of phenomena such as we have observed in cases which are similar, while hypothesis supposes something of a different kind from what we have directly observed, and frequently something which it would be impossible for us to observe directly. (Peirce, 1932, vol.2, p.385)

Abductive inference and the logic of discovery

Concerning the role of abduction in scientific inquiry, Peirce most consistently emphasized its operation in the initial creation of hypotheses and his ideas have been seen as providing the basis for a logic of scientific discovery (see also Hanson, 1958). To take one example, in a useful analysis of the arguments both for and against the possibility of such a logic, Curd (1980) attempts to resolve the debate by analyzing three major positions: the hypothetico-deductive account championed by Popper and Hempel; the inductive-probability account attributed to Reichenbach and Salmon; and the abductive inference account offered by Peirce. To begin, Curd argues it is necessary to distinguish between two fundamentally different components of a logic of discovery that are typically conflated: a logic of prior assessment and a logic of theory generation. According to Curd, a logic of prior assessment deals with “. . . the methodological appraisal of hypotheses after they have been generated but before they have been tested” (Curd, 1980, p.203). Such prior assessments can be of two kinds: a logic of probability, namely probability judgments of the truthfulness of a

hypothesis either in isolation or relative to competing hypotheses; and a logic of pursuit, which offers assessments of the pursuit-worthiness of a hypothesis or competing hypotheses based on more practical questions such as whether it makes sense to work on a particular hypothesis given certain constraints operating (e.g., time, money, resources, and the current state of knowledge), and which hypothesis of a range of possibilities recommends itself for detailed examination and testing (Curd, 1980).

The second component of a logic of discovery, namely a logic of theory generation, differs from a logic of prior assessment in that it is concerned with the initial creation of theories. In terms of possible characterizations of a logic of theory generation, Curd (1980) offers three alternatives: the production of a procedure or algorithm for generating meaningful hypotheses; historical narratives of periods of theory generation in the history of science; and a rational reconstruction of the theory generation process that includes both a classification and accompanying justification of the inferential strategies involved (Figure 3.1).

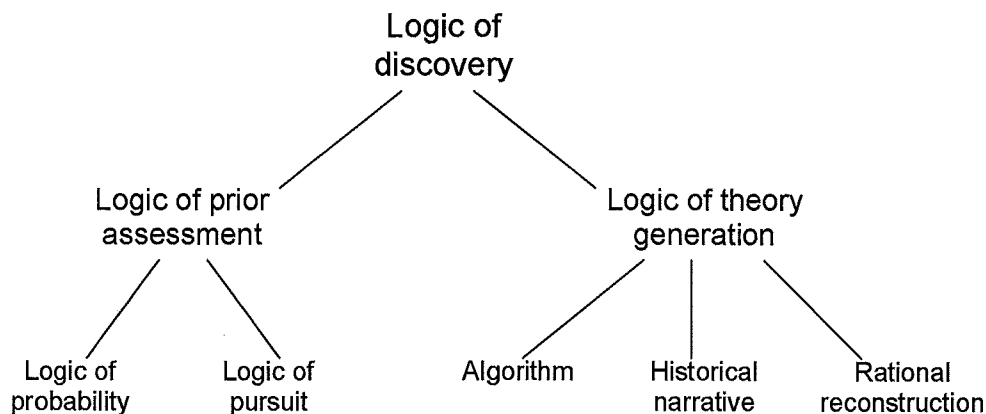


Figure 3.1 Curd's (1980) classification of the components of a logic of discovery.

Having made the above distinctions, Curd draws on Peirce's insights concerning abductive inference to argue that none of the objections traditionally raised against the possibility of a logic of discovery count against a logic of prior assessment understood

in terms of pursuit-worthiness or a logic of theory generation conceived as a rational reconstruction, and it is these senses of a logic of discovery that Peirce was committed to in his explication of abduction. Moreover, Curd suggests that these two conceptions are in fact reconcilable because the logic of prior assessment formulated as a logic of pursuit supplies the justification for the inferential strategies involved in a rational reconstruction. Therefore, Curd employs Peirce's broad ideas about abduction to argue that it is the logic of pursuit that is the key to constructing a logic of discovery (Curd, 1980).

*"Good Science is Abductive, Not Hypothetico-Deductive"*⁵

While Peirce's general treatment of abduction is most commonly referenced in connection with attempts to construct a logic of discovery, more recent investigations that extend his ideas to include detailed accounts of abductive inferential strategies, suggest the need to recognize a wider role for abduction in scientific inference. To take one example, Thagard (1988) has developed a detailed computational model of problem solving and discovery, which suggests that abductive reasoning is central to both the discovery and justification of scientific theories. Rejecting the traditional division between these two contexts of inquiry (e.g., Reichenbach, 1938), Thagard draws on computational analyses of the inferences involved to show that abduction is a key mechanism in the discovery of hypotheses as well as being an important element in their justification. Moreover, through the development of this computational model (named *PI* for "processes of induction"), and subsequent analyses of its performance on a range of different problem solving tasks, Thagard argues that there are a number of different ways in which hypotheses can be abductively obtained that have been overlooked by a traditional focus on 'simple abduction' or the formation of "hypotheses about individual objects" (Thagard, 1988, p.54). Specifically, he identifies four subspecies of abductive inference that are distinct from simple abduction, namely existential abduction, rule-forming abduction, analogical abduction, and inference to the best explanation. According to Thagard

⁵ Rozeboom, W. W. (1997). Good science is abductive, not hypothetico-deductive. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 335-391). Hillsdale, NJ: Erlbaum.

(1988) all four kinds of abductive inference have been neglected in attempts to characterize scientific inference, yet all can be recognized as playing an important role in the inquiry process.

Further work by artificial intelligence researchers who capture the algorithmic nature of how we reason abductively in a range of problem solving situations, reinforces Thagard's call for abductive reasoning to be allocated a central place in scientific methodology. For example, Josephson and Josephson (1994) report a range of findings made in the course of designing and testing systems that can successfully perform explanatory reasoning tasks and argue that abduction constitutes a family of "reasonable and knowledge-producing inferences" (Josephson & Josephson, 1994, p.1) with broad application in science and everyday life. And in philosophy of science, Rozeboom (1997) surveys a range of recent work that has contributed to the revival of interest in abduction, including his own research, which details the operational specifics of reasoning from data regularities to explanations of those regularities (e.g., Rozeboom, 1972). On the basis of these investigations, he argues that the abductive model of rational belief change offers a superior alternative to both the Bayesian and received hypothetico-deductive views of scientific method, and should be installed as "... our premier account of scientific reasoning" (Rozeboom, 1990, p.555).

Collectively, this body of recent work that investigates how we reason abductively has built on the general ideas formulated by Peirce to produce computationally rigorous and operationally specific guidelines for explanatory reasoning. In doing so, moreover, this research has highlighted the need to incorporate a central role for abductive inference in any adequate account of scientific method. Attending to the actual reasoning strategies deployed by scientists in their theory formation efforts, together with mounting evidence that abductive inferences offer the only viable means of reasoning successfully from phenomena to theory, suggests a focus for attempts to develop a theory of scientific method that surmounts the problems inherent in inductive and hypothetico-deductive accounts. It is time then, to turn to a theory of scientific method that attempts just this.

3.1.3 Abductive method: An overview

Against the background of a renewed philosophical interest in the empirical base of science and a growing support for abductive reasoning to occupy a central place in scientific methodology, a recently developed abductive account suggests itself as a promising candidate for a general theory of scientific method. In contrast to both inductive and hypothetico-deductive accounts, this theory, advanced by Brian Haig (Haig, 2002), takes scientific inquiry to be a problem-focused endeavour that is centrally concerned with the detection and explanation of empirical phenomena. According to the abductive theory, inquiry proceeds through a number of phases:

- Guided by an evolving problem that directs and constrains inquiry, data are collected and analysed in order to detect robust empirical regularities or phenomena.
- Once these phenomena have been reliably detected, they are explained by generating ideas about the mechanisms that are producing them. This abductive process involves reasoning back from the phenomena to the causal mechanisms underlying the phenomena in order to generate a rudimentary explanatory theory of those phenomena.
- Following its generation, the theory is developed by employing one or more modelling strategies that serve to elaborate on the nature of the causal mechanisms involved.
- When a theory is well developed, it is evaluated against competing theories on a range of empirical and super-empirical criteria that emphasize the explanatory worth of theories (Haig, 2002).

Theory construction is therefore conceived as an extended reasoning process in abductive method in which phenomena provide the focus for inquiry and abduction assumes a prominent role in all three contexts of theory generation, development, and evaluation. The main inferential move in abductive method from phenomena to theory is depicted in Figure 3.2:

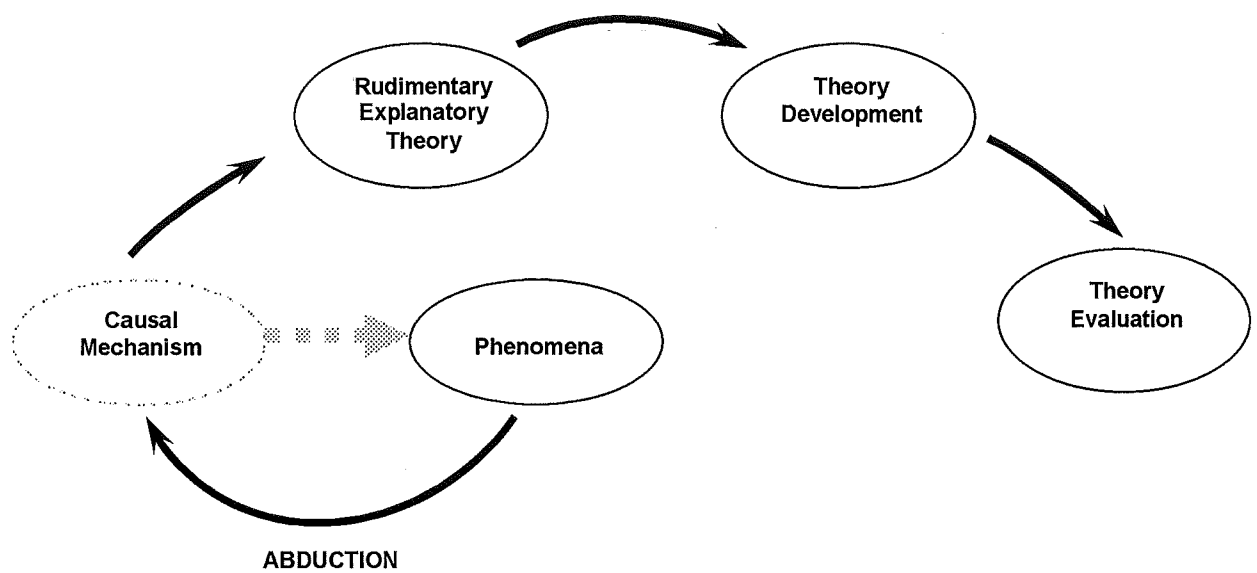


Figure 3.2 The explanatory move from phenomena to theory in the abductive method.

By recasting scientific method broadly in terms of the detection and explanation of empirical phenomena, the abductive theory is well placed to overcome the limitations of inductivism and hypothetico-deductivism highlighted in Chapter 2. Firstly, by codifying how these twin objectives guide scientific research, the abductive theory necessarily covers a broad range of investigative tasks. The result is a wholistic conception of method that includes the formulation of the research problem and the collection and analysis of data to facilitate phenomena detection, as well as the extended process of theory construction that incorporates generation, development, and appraisal dimensions.

Secondly, the abductive theory doesn't just provide a more complete framework for understanding scientific method. Instead, it embodies a significant reinterpretation of the 'theory-evidence' relationship that carves up their respective roles and responsibilities in a fundamentally different way. Whereas inductivism stresses the primacy of observed facts and hypothetico-deductivism emphasizes the primacy of theory, the abductive theory reinterprets the empirical base of science in terms of

phenomena and accords equal weighting to both their detection and theoretical explanation. According to abductive method, scientists construct theories in order to explain phenomena. The development of new knowledge is therefore grounded in, and constrained by, an initial search for robust empirical regularities. In turn, the theorizing process that follows goes beyond a description of the phenomena to abductively infer the existence of causal mechanisms responsible for their occurrence. Therefore, rather than the foundational generation advocated by naïve inductivism, or the foundational justification touted by hypothetico-deductivism, the abductive account codifies the methodological processes involved in moving from phenomena to theory to advance a form of what I will call '*empirically grounded explanation*'. The following sections clarify and elaborate this idea.

Thirdly, in its role as a general theory of scientific method, the abductive theory provides a broad conceptual framework that attempts to capture how scientists typically develop new knowledge. In contrast to inductive and hypothetico-deductive accounts, the abductive theory does not advocate a simplistic algorithm for knowledge production. Rather, it provides an integrative super-structure highlighting the multiple phases of inquiry, that frames and informs more specific methods and procedures. The framework or super-structure explains and regulates the knowledge construction process as a whole, while the specific research methods give the theory its operational force in each of the phases of inquiry that together comprise the general framework.

With a brief overview of the abductive theory of scientific method in place, the following sections look to provide a more detailed characterization of the framework. In undertaking this characterization, my aim is twofold: firstly, to clarify how the multiple interacting contexts of inquiry form a cohesive whole; and secondly, to show how this particular construal of the interplay between the 'mind' and the 'world' provides us with a workable general theory of scientific method.⁶

⁶ The following depiction of abductive method draws heavily on Haig (2002). For earlier formulations of the abductive theory see Haig (1987, 1996); for applications of the theory to clinical assessment see Ward and Haig (1997) and Ward, Vertue and Haig (1999).

3.2 The abductive framework

3.2.1 Research problems

In an attempt to improve on the strictures of inductivism and hypothetico-deductivism, Haig (2002) begins by emphasizing the importance of understanding method within the context of problem solving. While the selection and formulation of research problems is an integral part of scientific inquiry, neither of the orthodox theories of scientific method examined in the previous chapter takes the idea of a research problem seriously. In fact, this neglect of problems is arguably a contributing factor to the inability of inductivism and hypothetico-deductivism (standardly conceived) to adequately explain how science progresses. For example, an inductive account instructs researchers to gather data in a free and undirected manner but offers no guidance concerning which facts are relevant and therefore is incapable of showing how an investigation could get underway. A standard hypothetico-deductive account also ignores research problems by initiating inquiry paradoxically with theories or problem *solutions* (Nickles, 1981). However, in doing so, hypothetico-deductivism effectively confines its attention to the final phase of inquiry and fails to demonstrate how scientists actually reach this endpoint.

In contrast, Haig (2002) explicitly acknowledges the need for a problem-oriented conception of inquiry and incorporates a “constraint inclusion” account of problems within the abductive theory (Nickles, 1981; Haig, 1987). In brief, this account defines a problem as comprising all the constraints required for its solution as well as the demand that the solution be found. According to this characterization, research problems are not divorced from relevant background theory or from constraints imposed by heuristics, principles and rules. Rather, these constraints are included in the research problem itself and actually define what the problem is by determining its structure. Moreover by including the basic demand that they be solved, problems are necessarily connected to the goals of the research programme in which they arise. Adopting this account of problems then, not only indicates that problems sit at the heart of inquiry, but also reveals how they provide an inherent road map for the inquiry process. Specifically, given that articulating the research problem involves

articulating the constraints on what would count as an admissible solution, the problem effectively regulates inquiry by directing progress towards its own solution. As both Haig (1987) and Nickles (1981) remark, the constraint inclusion account of problems reveals that stating the problem is literally half the solution!

Incorporating this account of problems within the abductive theory, Haig contends that a scientific investigation is often launched by a poorly structured problem that evolves through the course of inquiry. From a constraint inclusion perspective, a poorly structured problem is one that lacks many of the constraints that are required to solve it. Building in these constraints to get a better idea of the nature of the problem (and by implication its solution) will generally occupy an extended period of research time. Because of this, Haig dismisses the idea that a problems component is merely the first phase of an investigation. Instead, he suggests that the problem extends right through the inquiry process, directing research by pointing the way to its own solution and therefore determining the parameters on scientists' decision making as the investigation moves through the contexts of phenomena detection and theory construction that comprise the abductive framework. Therefore, by incorporating a constraint inclusion account of problems, the abductive theory is able to both explain how inquiry proceeds and at the same time harness the regulative power of problems to effectively govern the investigative process (Haig, 2002).

3.2.2 Phenomena detection

Under the guidance of a developing problem that regulates and directs inquiry by imposing a number of constraints on the knowledge construction process, the abductive theory turns to the collection and analysis of data. In contrast to hypothetico-deductivism, which begins with a theory or hypothesis, abductive method, in accord with insights provided by the new experimentalists, demands that empirical facts be given precedence and emphasizes the role that these facts play in knowledge generation not merely its validation. However, unlike naïve inductivism, which fails to provide a tenable account of this generative process, the abductive theory outlines a defensible strategy for grounding scientific explanations in the world. Firstly, the initial collection and analysis of data is explicitly directed by an

evolving research problem, thereby overcoming the difficulties encountered by inductivism concerning which data should be collected and analysed. Secondly, the abductive theory endorses the fundamental distinction highlighted earlier between data and phenomena (Bogen & Woodward, 1988, 1992; Woodward, 1989, 2000):

Phenomena are relatively stable, recurrent general features of the world that we seek to explain. By contrast, data are idiosyncratic, ephemeral, and pliable and serve as observable evidence for phenomena. Phenomena comprise a varied ontological bag that includes objects, states, processes, events, and other features that are hard to classify. It is, therefore, more useful to characterize phenomena in terms of their *role* in relation to explanation and prediction. Phenomena, not data, are the proper objects of scientific explanation; it is phenomena that give scientific explanations their point; and it is the generality and stability of phenomena that make them the appropriate focus of theory construction. (Haig, 2002, p.4)

While both inductivism and hypothetico-deductivism ignore the task of phenomena detection, we saw in Section 3.1.1 that it is necessary for a satisfactory account of scientific method to adequately distinguish data and phenomena, and recognize the different roles that each play in scientific inquiry. Accordingly, the abductive theory takes phenomena, rather than data, to provide the focus for inquiry. Phenomena are what scientists attempt to construct systematic explanations about and in this sense comprise the most important constraint on directing the inquiry process. Because of this role in theory generation, Haig argues that it is vital that the phenomena in question are *robust* – that they are genuine effects and not merely artifacts of a particular experimental setup or statistical technique.⁷ Given this demand, and given the inherent difficulties involved in reliably establishing the existence of a genuine phenomenon highlighted for example by the new experimentalists' narratives (e.g.,

⁷ Phenomena can also be detected in non-experimental contexts. To take one example, meta-analysis comprises a bundle of data analytic techniques that is widely employed in psychological research to establish claims about phenomena or, as they are often termed, 'effects'.

Franklin, 1990), the abductive theory advocates a multi-stage model of data analysis that includes:

- *initial examination of the data* to determine its quality and structure as a means of assessing its suitability for further analyses (Chatfield, 1985);
- *exploratory data analysis*, utilising multiple forms of description and display in an effort to detect data patterns (Tukey, 1980);
- *confirmatory data analysis 1: close or internal replications* using computer intensive resampling techniques such as the jackknife, the bootstrap, and cross validation (Efron & Gong, 1983), to provide an initial confirmation or check on the reliability of the data patterns;
- *confirmatory data analysis 2: constructive or external replications* using different data sets and different measurement techniques to determine the reproducibility and therefore generalizability of the results. Successful constructive replications, which systematically vary the conditions under which the data patterns are produced, go beyond the initial confirmation provided by internal replications and are essential for justifying claims about the existence of phenomena (Haig, 2002).

Haig argues that all four stages of data analysis are typically required for phenomena detection. That is, he suggests it is only by sifting and resifting the data in this way and employing multiple means of establishing the nature and existence of phenomena that researchers can hope to reliably produce empirical facts worthy of scientific explanation.⁸

⁸ It is worth noting that Haig's acknowledgement of the need for an initial examination of data to assess its quality that is distinct from later phases of confirmatory data analysis where the concern is establishing claims about phenomena, is in accord with Woodward's (1989) contrast between data and phenomena in terms of the notions of error that are applicable to each. Specifically, Woodward points out that 'error' in the case of data relates to simple perceptual or recording mistakes, for example "... misreading a dial or transposing digits when a number is entered into a laboratory notebook" (Woodward, 1989, p.394). In contrast, 'error' in relation to phenomena involves much more complicated and subtle kinds of problems such as "... failure to adequately control for various background and confounding factors or mistakes in statistical analysis or in procedures for data reduction (Woodward, 1989, p.394). On the model of data analysis outlined above, when researchers undertake a preliminary screening of their data during the initial phase of analysis, they are concerned with identifying errors of the first sort. In contrast, it is only during subsequent confirmatory data-analysis phases, that the concern is with errors of the second sort, i.e., whether one is detecting a real effect or an artifact of the experimental setup or data analysis/detection procedures employed.

3.2.3 Theory generation

According to the abductive theory, it is necessary to distinguish between data and phenomena and deploy multiple data analyses in order to unearth facts adequate for science. Once obtained, these facts or descriptive claims about phenomena will require systematic explanation. For orthodox inductivism, theory generation is an inductive reasoning process in which theories are somehow mechanically derived from a secure observational base. For hypothetico-deductivism, theory generation is amethodological, involving free use of the imagination. For abductive method, empirical phenomena provide the stimulus for theory generation, and the explanatory move from phenomena to theory is achieved through abductive reasoning (Haig, 2002).

Specifically, Haig endorses Peirce's insights about the creative inference involved in theory formation and argues that the generation of theory is typically a process of reasoning back from puzzling phenomena to an explanation of the causes underlying the phenomena. Reworking Peirce's characterization to include the recognition that theories explain phenomena rather than observed data, Haig presents the moves involved in abductive inference as follows:

[S]ome observations (phenomena) are encountered which are surprising because they do not follow from any accepted hypothesis; we come to notice that those observations (phenomena) would follow as a matter of course from the truth of a new hypothesis in conjunction with accepted auxiliary claims; we therefore conclude that the new hypothesis is plausible and thus deserves to be seriously entertained and further investigated. (Haig, 1996, p.286)

In connection with this characterization of abductive reasoning, Haig draws attention to a number of regulative principles that constrain abductive inferences, and thereby ensure that scientists generate theories that provide the most *plausible* explanations of the phenomena. Discovery is not conceived as a process in which any novel hypothesis will be entertained, but as one in which only the most plausible hypothesis,

determined by reasoning abductively within a context of regulative constraints, is considered worthy of pursuit (Haig, 2002).

Moreover, Haig endorses Thagard's (1988) differentiation of Peircean abduction into a number of different subspecies, and argues that three of these subspecies are basic to the abductive account of scientific method and can be distinguished in terms of their respective roles in the framework. In particular, existential abduction, which postulates the existence of previously unknown objects, is located in the context of theory generation.⁹ Analogical abduction, which employs past cases of hypothesis formation to produce new hypotheses, is situated in the context of theory development but can also function in the context of theory generation. Finally, inference to the best explanation, which involves the comparative appraisal of mature theories, occurs in the context of theory evaluation.

The suggestion that abduction is a valid means of generating knowledge, runs counter to the view endorsed by proponents of hypothetico-deductivism (e.g., Hempel, 1966; Popper, 1959) who argue that there is no logic to discovery. For example, Popper argues:

The initial stage, the act of conceiving or inventing a theory, seems to me neither to call for logical analysis nor to be susceptible to it. The question how it happens that a new idea occurs to a man – whether it is a musical theme, a dramatic conflict, or a scientific theory – may be of great interest to empirical psychology; but it is irrelevant to the logical analysis of scientific knowledge . . . My view of the matter . . . is that there is no such thing as a logical method of having new ideas, or a logical reconstruction of this process (Popper, 1959, pp.31-32).

I argued in Chapter 2 that one of the predominant reasons why hypothetico-deductivism fails to provide a methodological account of theory generation, viewing it

⁹ Exploratory factor analysis is one example of a generative method used in psychological research that employs existential abduction to reason from correlational data to factorial proto theories via the principle of the common cause.

as an activity beyond rational characterization, is that it operates with an overly restrictive model of scientific reasoning centred on deductive logic. In contrast, the abductive theory rejects such a narrow construal of rationality and is therefore able to characterize the supposedly indescribable ‘creative leap’ as a “discursive reasoning complex” (Haig, 1987) that pivots on the natural human ability for abductive inference.

Having admitted a logic to discovery however, the abductive theory stops short of the assumption embodied in naïve inductivism that scientists could algorithmically generate true theories from data. Rather, it is envisaged that scientists’ natural talent for abductive inference, appropriately constrained by various regulative principles, is more likely to produce a number of plausible explanations of the phenomena that will require further investigation. Therefore, Haig, in accord with Curd’s (1980) interpretation of Peircean abduction outlined in Section 3.1.2, explicitly characterizes the logic embodied in abduction as a logic of “pursuit”, and emphasizes the need for a recognition and tolerance of pluralism in this generative phase of inquiry.

3.2.4 Theory development

In Chapter 2, we saw that the hypothetico-deductive account of scientific method not only places theory generation outside of science, but also ignores the methodological process of theory development. The fact that this received view begins with theory testing, only serves to encourage the assumption that theories are generated in a mature form. However, this assumption sustains the myth that discovery is an event rather than a process (Curd, 1980), and ignores the reality of scientific practice where new ideas typically require nurturing and maturation. In particular, theory development frequently involves extending knowledge of a theory’s proposed causal mechanisms. Critical to this development process, is the application and employment of models (Haig, 2002).

In a comprehensive account of the role of models in scientific inquiry, Harré (1976) reveals how a model as a representative device can enable scientists to develop their understanding of a theory’s underlying causal mechanisms. In Harré’s view, the

creative task involves inventing a plausible analogue of the causal mechanism that is responsible for the phenomenon. This is achieved by enlisting an appropriate source model, often via a process of analogical abduction. A number of constraints on the modelling process, such as the fact that the mechanism must behave analogously in relevant respects to the known source, and the need for the model to maintain a relation with the real processes and patterns of nature, serve to constrain analogical reasoning and encourage a developed formulation of the scientific theory (Harré, 1976).

While both inductivism and hypothetico-deductivism treat models simply as dispensable psychological aids, the abductive theory encourages the incorporation of a variety of different modelling strategies into the inquiry process and allocates them a genuine methodological role within its framework. Specifically, Haig (2002) identifies two modelling strategies that are important in scientific theory construction. Firstly, paramorphic models, which have a different source and subject, such as the widely applied thinking-as-computation model in cognitive science research, can be used to facilitate a detailed knowledge of a theory's proposed causal mechanisms. Secondly, homoeomorphic models, in which the source and subject are the same, such as causal modelling methods employing structural equation modelling procedures, can be effective in building up a picture of the wider causal network into which the proposed mechanisms enter. Haig (2002) argues that details of the causal mechanisms themselves and details of the causal network of which they form part are compatible kinds of causal knowledge, and both may be utilized during theory development.

3.2.5 Theory evaluation

Traditional approaches to theory evaluation, for example hypothetico-deductive and Bayesian accounts, largely confine their focus to the issue of empirical adequacy understood as a theory's predictive success. In contrast, the abductive theory endorses a flexible, multi-criterial perspective on theory evaluation. More particularly, Haig (2002) takes the evaluation of well-developed or mature theories to be largely a matter of establishing their explanatory worth through inference to the

best explanation (Harman, 1965). Inference to the best explanation involves accepting a theory when it is seen to provide a better explanation of the phenomena under investigation than its competitors do. As Haig (2002) remarks:

The phrase, *inference to the best explanation*, captures the basic idea that much of what we know about the world is based on considerations of explanatory worth and it involves the process of judging the best of competing explanatory theories. (Haig, 2002, p.8)

Critics of Harman's account have suggested that his idea is too vague to provide adequate guidance for theory appraisal. More recently however, Paul Thagard (e.g., Thagard, 1989, 1992) has provided researchers with a workable formulation of inference to the best explanation that involves making judgments of explanatory coherence. According to this formulation, which has been implemented in a computer program (ECHO) and has demonstrated widespread application, the evaluation of competing theoretical candidates is determined on the basis of three criteria: consilience or explanatory breadth, simplicity, and analogy. The criterion of explanatory breadth judges a theory to be more explanatorily coherent if it explains more facts than its competitors. The criterion of simplicity captures the idea that preference should be given to theories that require fewer ad hoc hypotheses. The third criterion, analogy, indicates that a theory demonstrates greater explanatory coherence if it is consistent with currently accepted theories that explain similar phenomena. Further discussion of Thagard's theory of explanatory coherence and its implementation in ECHO is given in Chapter 5.

Haig (2002) endorses Thagard's theory of explanatory coherence as a useful framework for assessing the explanatory worth of theories, but points out that it intentionally omits the criterion of predictive success from its evaluative machinery. He argues that while prediction is typically overemphasized in theory evaluation, and often at the expense of explanation (Haig & Durrant, 2000) the predictive success of a theory remains an important criterion of a theory's worth. Accordingly, there is a need to recognize that a theory's capacity for successful predictions will often feature in evaluative contexts, with its role dependent upon the nature of the theory under investigation. In addition, Haig draws attention to a number of other criteria that are

often utilized in scientific theory appraisal, namely initial plausibility, existential depth, and fertility, and suggests that all deserve consideration in a flexible model of theory evaluation that is adequate to the realities of scientific practice.

3.3 The abductive theory as a general theory of scientific method

In this chapter, I have presented an abductive theory of scientific method and suggested that it provides us with a workable theory of scientific inquiry. As an introduction to the abductive account, I began with an examination of recent work by the new experimentalists who argue for a renewed focus on the empirical base of science. Following Hacking's (1983) commentary I highlighted how, due largely to the prevailing philosophies of first positivism and then the theory-dependence movement, the importance of empirical phenomena to the development of scientific knowledge has been largely ignored by orthodox accounts of scientific method. I further suggested that a more adequate theory of inquiry needs to recognize the importance of codifying the methods by which scientists arrive at reliable empirical facts and, as part of this process, acknowledge that these facts will involve claims about phenomena rather than claims about observed data.

I then highlighted an increased interest by philosophers in abduction as an important species of scientific inference. Beginning with the work of Peirce (1931-58), I demonstrated that a focus on abductive inference offers a valid means of characterizing the supposedly indescribable creative move from phenomena to explanatory theory. I further outlined recent developments in artificial intelligence research on problem solving, which not only provide computationally rigorous and operationally specific guidelines for explanatory reasoning, but also indicate the need to recognize a central role for abduction in scientific inference, over and above the focus of orthodox accounts on induction and deduction.

Against this background, I presented Haig's abductive theory of scientific method (Haig, 2002). I argued that this theory incorporates insights from both these movements, and offers a promising alternative to orthodox accounts of scientific method.

Firstly, the abductive theory allocates phenomena a central role in its characterization of scientific method and gives equal methodological attention to both their detection and explanation. In order to adequately articulate these two major foci of scientific research, the abductive theory necessarily attends to much more of the inquiry process than either naïve inductivism or hypothetico-deductivism, and is therefore well placed to provide an informative perspective on scientific inquiry. In particular, the abductive theory offers methodologists a wholistic conception of method that incorporates a positive account of research problems, and recognizes the extended and often difficult nature of phenomena detection via multiple phases of data analysis, as well as highlighting the various contexts of theory construction that incorporate generation, development, and appraisal dimensions.

Secondly, I have attempted to show that the abductive theory does not merely furnish a more complete account of scientific method, but also embodies a significant reinterpretation of the relationship between 'theories' and 'evidence' that I have labelled 'empirically grounded explanation'. Specifically, on abductive method's characterization of the inquiry process, scientists construct theories in order to explain phenomena. The development of new knowledge is therefore grounded in, and constrained by, an initial search for robust empirical regularities. Moreover, the theorizing process that follows goes beyond a description of the phenomena detected to postulate via an abductive reasoning process the causal mechanisms that are producing the phenomena. Therefore, in contrast to the untenable foundational generation advanced by naïve inductivism and the unsatisfactory foundational justification touted by hypothetico-deductivism, I have argued that the abductive theory offers a defensible strategy for grounding scientific explanations in the world.

Finally, in its role as a general theory of scientific method, the abductive theory provides a broad conceptual framework that attempts to capture how scientists typically develop new knowledge. Unlike naïve inductivism, the abductive account

does not advocate a simplistic algorithm for truth production; neither does it advance a narrow prescription for evaluating theories in terms of their empirical consequences as promoted by the received hypothetico-deductive view. Rather, the abductive theory offers an integrative framework or super-structure that both frames and informs a variety of specific research methods. The superstructure explains and regulates the knowledge construction process as a whole, while the specific research methods give the theory its operational force in each of the contexts that comprise the general framework.

The abductive theory, then, offers a useful general theory of scientific method, but as Haig points out, it is not to be understood as an all purpose or universal account of method (Haig, 2002). A number of recent commentators on the nature of science (e.g., Chalmers, 1999; Nickles, 1990) have argued against the view that scientists in all disciplines employ a single method of inquiry. Drawing on contemporary analyses of the history of science, these commentators suggest that methods will necessarily be dependent on the particular subject matter under investigation and therefore are more appropriately conceptualized as domain-specific rather than domain-general in their application. Proponents of both the inductive and hypothetico-deductive accounts have been criticized for claiming that their respective account captures ‘the scientific method’, and for their failure to recognize the diverse modes of investigation actually operating in science. In contrast, the abductive theory, by providing a broad conceptual framework within which more specific research methods can be located and understood, has the flexibility to recognize the domain and content specificity of methods in science highlighted by these commentators, while at the same time providing an informative general perspective on the inquiry process.

In conclusion, I have argued that the abductive theory of scientific method provides an adequate third alternative to the orthodox accounts of method that were examined in Chapter 2. Specifically, by adopting a problem-oriented conception of inquiry and by giving attention to both the detection and explanation of empirical phenomena, the abductive account offers an informative portrayal of scientific inquiry that is not undermined by arguments for methodological pluralism in science. In the chapters that follow, I consider whether this abductive framework can also function as a source

model for the child-as-scientist analogy, and inform deliberations about the development of knowledge in childhood.

Chapter 4

From Source Model to Target: Mapping Relations between Scientists and Children

In Chapter 3, I presented an abductive theory of scientific method and argued that it offers researchers an informative general perspective on scientific inquiry. Having selected this theory as an appropriate source model for the analogy, the focus now turns to the issue of establishing meaningful relations between this account of scientific method and the knowledge construction efforts of young children.

With this aim in mind, Section 4.1 raises questions about the scope of Gopnik and Meltzoff's (1997) application of the theory to the development of knowledge in childhood, and suggests that their attempts to use the analogy to define a broad constructivist view of cognitive development do not facilitate the mapping of deep correspondences between scientists and children. I then outline an alternative interactionist perspective on development advanced by Jeffrey Elman and colleagues (Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996), and argue that it offers a more promising framework theory for cognitive developmental research. Section 4.2 subsequently looks to define a narrower role for my methodological formulation of the child-as-scientist analogy, and in doing so reveals some compelling parallels between the theory building strategies of scientists highlighted by abductive method and the problem solving activities of children detailed by Annette Karmiloff-Smith (Karmiloff-Smith, 1992). In Section 4.3, I conclude that enlisting the abductive framework and mapping relations from this source model to children's problem solving strategies uncovers deep correspondences between scientists and children that indicate the utility of this approach for investigating cognitive change.

4.1 “Interactions all the way down”: An alternative perspective on development

4.1.1 The story so far

In the Introduction to this thesis, I surveyed a range of positions taken by researchers on the question of whether the child-as-scientist analogy has utility as a model for cognitive developmental research. For those advocating the analogy as a source of useful ideas, I pointed to varying interpretations of what this might mean in terms of actual similarities or commonalities existing between children and scientists. These included the proposal that children’s everyday knowledge inheres in framework theories, claims for structural correspondences between knowledge systems and the ways in which these systems are restructured over time, and process-oriented analyses of the interplay between theories, heuristics, and data in effective problem solving. Similarly, for those researchers opposing the analogy, I highlighted a number of different reasons for their rejection of child-scientist parallels. These ranged from empirical claims for significant dis-analogies between children and scientists, to a priori arguments concerning the improbability of common mechanisms from the standpoints of developmental theory and social-constructionist analyses of the history of science.

Following this overview, I suggested that attempts to clarify the issue of child-scientists parallels would benefit from current investigations of the nature of analogy itself, and the ways in which effective analogical reasoning promotes theoretical innovation in science. Having identified the basic principles of analogical thinking common to many philosophical and cognitive science accounts, I subsequently outlined a multiconstraint theory of analogy that characterizes analogical thinking in terms of the interplay of multiple competing constraints and identifies a range of levels at which mappings can occur. Adopting this multiconstraint theory as a template for the child-as-scientist analogy, I suggested that a productive formulation of child-scientist parallels will necessarily promote system mappings of higher order relations between children and scientists, and will achieve these mappings by applying the constraints of similarity, structure, and purpose.

Following these methodological provisions for developing a productive scientific analogy, Chapter 1 turned to the most recent and detailed formulation of child-scientist parallels provided by Gopnik and Meltzoff's (1997) theory theory. After considering the major tenets of their position and highlighting some of the criticisms that have been levelled against it, I argued that these researchers fail to demonstrate convincingly that the cognitive processes of theory construction and revision are common to scientists and children. Moreover, I suggested that the theory theory is unable to offer any substantive ideas about the mechanisms responsible for cognitive development, and this is largely due to its dependence on an inappropriate source model of scientific inquiry.

With this diagnosis in hand, Chapter 2 focused on the task of selecting a more appropriate source model for the child-as-scientist analogy. Working from an evolutionary epistemological standpoint, I pointed out that due consideration of our evolutionary origins strongly supports a methods perspective on science and leads us to expect fundamental continuities, rather than discontinuities, between everyday cognition and scientific inquiry methods. Further, I showed how Gopnik and Meltzoff's focus on theory-evidence relations to explain conceptual and linguistic developments requires a methods-centred view of science. I therefore concluded that if the aim is to develop the child-as-scientist analogy beyond its current heuristic status, then a methodological reformulation must take place.

Having considered and discounted the two major orthodox theories of scientific method as potential source models for the analogy, Chapter 3 presented an alternative abductive theory of scientific method. I argued that this theory offers an informative general perspective on the process of knowledge construction in science, and therefore has the potential to inform deliberations about cognitive change in childhood. With this source model in place, the following chapters develop my proposed methodological model of cognitive change. More specifically, the focus of this chapter is to establish the initial plausibility of the claim that the abductive framework offers a source of ideas about children's knowledge construction strategies. In order to do this, I return to the broader issue of how best to characterize cognitive development, and to the major criticism of the theory theory highlighted in earlier chapters, namely its failure to offer any substantive ideas about the processes

through which children's theories are developed. Using this criticism as a starting point, I consider an interactionist account of cognitive development as an alternative general framework for investigating developmental change, before going on to define a narrower role for my proposed model of child-scientist parallels that focuses on one level of organism-environment interactions in particular.

4.1.2 Navigating between the poles of nativism and empiricism: Attempts to chart a viable "radical middle"

In earlier chapters, I indicated that the most serious charge against the theory theory is its failure to explain how theory development occurs in the young child. By recruiting Kuhn's description of scientific change, and couching conceptual and semantic developments in terms of theory-evidence relations, Gopnik and Meltzoff (1997) suggest that the burden of explanation for cognitive development lies with the same mechanisms that are responsible for development in science. Yet detailing the ways in which evidence motivates children to change their theories, in other words, explaining how experience *interacts* with existing representations of the world to produce new representations, remains profoundly mysterious on the theory theory model.

Another way in which to highlight these concerns is to consider the theory theory's claim on what has become known as the "radical middle". In a recent review article (Newcombe, 1998; see also Newcombe & Learmonth, 1999), Nora Newcombe identifies an emerging trend amongst cognitive developmentalists to eschew the extremes of nativist and empiricist approaches to learning and cognition:

In the past several years . . . there has been a mounting backlash against nativist dogma, as well as a continuing suspicion about simple empiricism. There is a feeling of excitement and competition abroad, as investigators vie to define a view of cognitive development neither radically nativist nor radically empiricist – a possibility I have heard called the 'radical middle'. (Newcombe, 1998, p.210)

As Newcombe points out, the trend is not towards replacing a radical nativist position with an equally radical empiricist one. Nor is it an attempt to resurrect Piaget's particular version of constructivism. Rather, the aim of these recent theoretical endeavours is to develop a new kind of developmental theory, in which the traditional dichotomies of nature versus nurture and continuity versus change are discarded in favour of a genuinely *interactional* account of development (Figure 4.1).

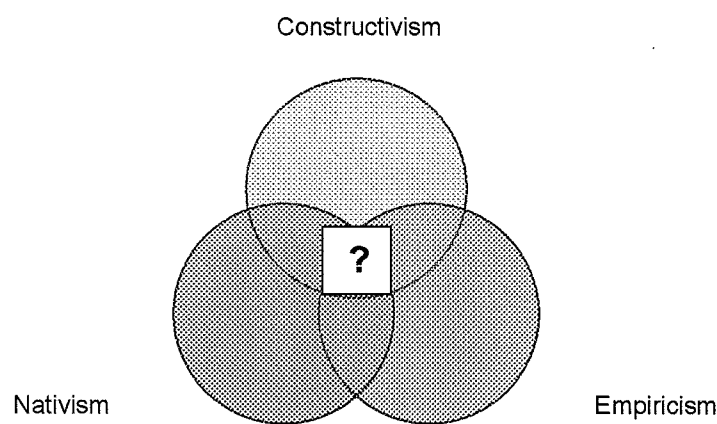


Figure 4.1 Claims on the “radical middle” position: attempts to define a new theory of cognitive development at the ‘intersection’ of nativism, empiricism and constructivism.

Working from Newcombe's characterization, Gopnik and Meltzoff's (1997) 'cognitive development as theory change' approach represents one attempt to claim this radical middle ground.¹ Specifically, by adopting a model of revolutionary scientific change and by proposing that the key to explaining knowledge development lies in the interactions between theories and evidence in science, Gopnik and Meltzoff draw on the analogy with science to lay the groundwork for a viable constructivist

¹ In addition to Gopnik & Meltzoff (1997), Newcombe identifies Karmiloff-Smith (1992), Thelen & Smith (1994), Siegler (1996), and Elman et al. (1996) with this trend.

theory of cognitive development.² However, in line with the criticisms highlighted in Chapters 1 and 2, the theory theory can be seen to fail in its bid to define the ‘radical middle’ on at least two counts.

Firstly, while Gopnik and Meltzoff (1997) argue that thinking of cognitive development as theory change offers a workable interactionist framework for cognitive developmental research, they demonstrate an unwillingness to relinquish their commitment to a strong form of representational nativism. As they say, “. . . our claim that there are innate theories . . . is in many ways as powerfully nativist a position as one could have” (Gopnik & Meltzoff, 1997, p.221). Yet by promoting the idea of “elaborate, rich, representational structure from birth” (p.220), Gopnik and Meltzoff deny themselves the opportunity to rid their model of the ‘innate’ versus ‘learned’ dichotomy that has proved such a stumbling block for traditional accounts. The result is an uneasy alliance between a rich representational endowment on the one hand and a revisionist empiricism on the other. Despite Gopnik and Meltzoff’s claims to the contrary, the nature of this alliance suggests that the theory theory is less a viable new interactionist position on development than a compromise between nativist and empiricist views.

Secondly, in addition to retaining a strong version of nativism, Gopnik and Meltzoff’s theory theory offers few details about the ways in which intrinsic structure and a structured environment interact to shape cognition and behaviour. As other contributors to the nativist-empiricist dialogue have pointed out:

. . . the problem is not so much that we do not know what the sources of knowledge are. The problem is rather in knowing how these sources combine and interact. The answer is not Nature *or* Nurture; it’s Nature *and* Nurture. But to say that is to trade one platitude for another; what is necessary is to understand the nature of that interaction. (Elman et al., 1996, p.357)

² In addition to Gopnik and Meltzoff’s (1997) presentation of the theory theory, further support for this interpretation is found in Gopnik’s (1996a) commentary on the post-Piagetian era, in which she promotes the theory theory as “[t]he most influential contemporary constructivist theory” in cognitive developmental research (Gopnik, 1996a, p.221).

By redefining children's conceptual structures as theories and giving experience an evidential role in a theory revision process, Gopnik and Meltzoff try to engage a model of scientific change in an attempt to illuminate this interactive process. However, the paucity of ideas about how such interactions, couched in theory-evidence terms, actually occur, substantially weakens the theory theory's claim to have shown how children derive new representations of the world from their experience. These problems with the theory theory, then, suggest the need to consider alternative ways of defining development that are neither radically nativist nor radically empiricist, and that offer some substantive ideas about the reciprocal actions of organism and environment in producing developmental change. With this aim in mind, the following section outlines a broad perspective on development that makes a strong bid to chart such a 'radical middle' course. According to the proponents of this approach, the key to capturing the nature of development lies in the recognition that it is 'interactions all the way down'.

4.1.3 Development as a truly interactive process

While Gopnik and Meltzoff (1997) look to the analogy with science to provide answers about the knowledge acquisition process in the young child, an alternative perspective on development suggests that such answers are far more likely to arise from the intersection of developmental neurobiology and connectionist modelling. In a recent collaboration (Elman et al., 1996), Jeffrey Elman and colleagues put forward their views on development, the nature/nurture controversy, and the question of innateness, from the standpoint of a biologically inspired connectionism. They argue that dramatic advances in the neurosciences coupled with the conceptual and computational tools provided by neural network or connectionist models, offer a new and powerful framework for investigating developmental change:

Two recent developments . . . suggest that the view of development as an interactive process is indeed the correct one, and that a formal theory of emergent form may be within our grasp. The first development is the extraordinary progress that has been made in the neurosciences. The second has been the renaissance of a computational framework which is

particularly well suited to exploring these new biological discoveries via modeling. (Elman et al., 1996, p.2)

In particular, the authors argue that this framework encourages a truly interactive perspective on development, according to which:

- (1) the responsibility for change is far more evenly distributed between a structured environment and innate predispositions than with Gopnik and Meltzoff's theory;
- (2) conceptual and linguistic achievements are reconceived as *emergent* products of a dynamic interactive process.

The following sections elaborate on this sketch by highlighting the basic commitments underlying the authors' approach, before moving on to consider some of the key insights offered by this theoretical framework for developmental psychology.

Basic commitments

Biology and connectionism: A natural synthesis Elman et al. (1996) begin their presentation by considering the problem of change. They argue that while an interactionist view of development is an attractive alternative to the extremes of nativism and empiricism, traditional formulations have been hampered by the lack of details about the nature of gene-environment interactions.³ According to Elman et al., however, recent advances in two fields not only endorse an interactionist view but also have the potential to advance it by providing the basis for a formal theory of emergent form. In the neurosciences, Elman et al. highlight recent findings regarding the complexity of genetic functioning and the indirectness of genes' effects on the emerging phenotype that overturn the assumption of a static genetic blueprint for biological form and behaviour. In addition to this research on the genetic basis of behaviour, they highlight current data from a range of sources indicating that the brain

³ For a classic presentation of the interactionist position, see Waddington (1975).

is initially equipotent (or at least multipotent) at the cortical level.⁴ According to Elman et al., these findings undermine claims for fixed, immutable forms of neural organization, and instead support a model in which patterns of regional specialization are progressively built up over time.

In combination with these recent advances in the neurosciences, Elman et al. draw attention to the field of computational modelling, where there has been a resurgence of interest in neural network or connectionist models. Elman et al. argue that these models are particularly pertinent to the concerns of developmental researchers because they offer concrete demonstrations of how the application of simple learning algorithms operating on local information can produce global behaviours (see also Plunkett, Karmiloff-Smith, Bates, Elman & Johnson, 1997). In particular, such demonstrations have forced researchers to revisit assumptions about what can be learned as opposed to what is prespecified, and to recognize that far more structure is latent in the environment and capable of being abstracted by basic learning algorithms than previously imagined. Elman et al. propose that when taken together, these advances in neural network modelling and in the neurosciences are mutually constraining and serve to form the basis of a powerful framework for rethinking fundamental issues of developmental change.

The role of interaction in development Elman et al. stress throughout their defence that adopting a bio-connectionist framework reveals the fundamental dependence of development on multi-level interactions. They highlight research in molecular genetics that underscores the need to investigate the workings of genes in concert to build an accurate picture of genetic functioning. Similarly, they point out that at higher levels of organization, recognizing the dependence of processes such as cell differentiation and tissue formation in regulatory systems on multiple interactions, is crucial to understanding how complexity is progressively built in to a system over time. Elman and colleagues argue that while developmentalists have long recognized the importance of interactions to development, the problem has been how to formalize this relationship. They propose that the ability of connectionist models to capture the

⁴ See Johnson (1998) for a review of neural development that emphasizes the plasticity of the developing brain.

dynamics of interactions between pre-specified biases and a learning environment in an artificial system over time, offers a potential solution to this problem. Moreover, the striking resemblance between this process of error reduction in an artificial neural network and earlier depictions of the developmental process in natural systems (e.g., Waddington, 1975), is taken as further evidence of the utility of connectionism to formalize the interactional nature of development.

Clarifying innateness A key part of the authors' argument for the utility of their biologically oriented connectionism concerns the implications of this framework for understanding innateness. They suggest that although few researchers would propose a simple one-to-one relationship between genes and behaviour, far greater clarity needs to be injected into current discussions of innate properties in the developmental and cognitive science literature. Moreover, Elman et al. point out that very little attention is given to issues of biological plausibility in these discussions. That is, they argue "... [t]he problem with current nativist theories is that they offer no serious account of what it might mean in biological terms for something to be innate" (Elman et al., 1996, p.48). Finally, Elman et al. emphasize that despite the current dominance of nativist approaches to the development of language and cognition, there has been little investigation of the potential variety of ways in which behaviours could be innate.

In an attempt to address these issues, Elman et al. begin by adopting a working definition of 'innate' that refers to features of brain structure, cognition, or overt behaviour that result from interactions that occur within the organism, as opposed to interactions between the organism and its external environment (see also Johnson & Morton, 1991; Johnson, 1998). On this view, as the authors make clear, the term 'innate' is not equivalent to 'coded in the genes'. With this working definition in place, Elman et al. proceed to construct a taxonomy of ways in which properties can be usefully classified as innate that attempts to identify the range of possible constraints on development. The result is a three-layered taxonomy that classifies constraints at the levels of representation, architecture, and timing.

To begin, positing the existence of representational constraints on development suggests that the representations that underlie knowledge are themselves innate or

'hard-wired'. This claim for *representational nativism* is common to accounts of development that argue for innate domain-specific knowledge systems, irrespective of whether these systems are conceptualized as modules in the Fodorian sense (e.g., Leslie, 1992; Spelke, 1994), or as indicated earlier, innate theories (Gopnik & Meltzoff, 1997). Elman et al. identify representational nativism as the strongest level of claims about innateness because of the direct and unmediated relationship assumed to obtain between the innate representations and the resulting knowledge and behaviour. However, they argue that given neuroscientific evidence of substantial cortical plasticity, innate representations – understood as “fine-grained patterns of synaptic connectivity at the cortical level” (p.25) – are likely to be relatively rare. Accordingly, they argue that there is a need to identify alternative and more biologically viable ways in which innate constraints could operate in development.

While Elman et al. argue that evidence supporting the tenets of representational nativism is unlikely to be forthcoming, they suggest that claims for constraints at the levels of architecture and timing demonstrate far greater promise. Architectural constraints, which they subdivide into unit-based (e.g., specification of neuron types), local (e.g., number of layers, packing density of cells), and global (e.g., connections between regions of the brain, input/output channels), raise the possibility of critical pressures operating at the level of the structure of the network or subparts of the network rather than at the level of representations. Elman et al. propose that such pressures, while less direct and specific in their effects, could nevertheless significantly constrain cognition and behaviour, for example, by determining that a specific cortical region receives a particular form of input that in turn dictates the type of representations developed by that region. Finally, the possibility of chronotopic constraints suggests that the actual timing of developmental sequences is a significant force in determining outcomes. That is, instead of solutions to problems being encoded in advance in the form of innate representations, such solutions could be rendered inevitable through the particular timing of input to the brain. Elman et al. argue that while behaviours will typically result from the operation of constraints at multiple interacting levels, further investigations of the possibilities suggested by both *architectural nativism* and particularly, *chronotopic nativism*, are likely to reveal powerful mechanisms for developmental change.

Key insights for developmental psychology

Having outlined the basic commitments underlying the bio-connectionist approach, it is useful to briefly consider some of the key insights for developmental psychology afforded by this theoretical framework. In other words, how does an account that rejects complex representational start states in favour of complex organism-environment interactions, and does so via the synthesis of biological and connectionist insights, offer researchers a viable alternative theory of cognitive development? Anticipating this question, Elman et al. identify the primary contributions of their approach as follows.

Innateness and domain-specificity: Distinguishing mechanism from content

To begin, constructing a taxonomy of innate constraints on development that identifies the multiple levels of representations, architectures, and timing, serves to clarify understanding of innateness in at least two fundamental respects. Firstly, this biologically informed taxonomy indicates the potential variety of ways in which a behaviour could be classified as innate, and in doing so challenges researchers to be far more precise in their discussions of innate properties. Secondly, this taxonomy marks an important distinction between innate *mechanisms* and innate *content*. Elman et al. stress that the trio of mechanisms they postulate is logically independent of the possible content domains over which these mechanisms operate. Moreover, this distinction is critical to the task of teasing apart issues of innateness from issues of domain-specificity, where ‘specificity’ may relate to any one of a number of different levels including tasks, behaviours, representations, processing mechanisms, and genes. By distinguishing mechanisms and content, Elman et al. highlight the inferential leap that separates claims about the specificity of certain abilities from claims about the nature of the mechanisms that subserve those abilities, and propose that in the case of higher-level cognition, most domain-specific outcomes are likely to result from the operation of domain-general mechanisms.

Non-linear change One of the key insights provided by connectionist models is that the mapping between overt behaviour and underlying mechanism is often non-linear. Elman et al. stress that contrary to assumptions underpinning much developmental research, qualitative changes in behaviour do not necessarily signal

qualitative changes in the mechanisms responsible for the behaviour. Instead, these models demonstrate that sudden dramatic effects in terms of the output of a system can be produced by small, incremental changes in internal processing. In the case of ontogenetic development, this suggests that apparent discontinuities in conceptual or linguistic understanding may not be the outcome of new mechanisms coming 'on-line' at certain points in development as has often been assumed, but instead reflect the continuous operation of a single mechanism over time.

Outcome and cause: Not a one-to-one relation In addition to showing that a single mechanism can be responsible for multiple behaviours, connectionist models can also illuminate the reverse case in which a single outcome or behaviour arises through the action of multiple interacting mechanisms. Further, Elman et al. point to instances where the same behavioural outcome can be produced in a number of different ways, as in the case of degraded performance in artificial neural networks. Precisely because they allow researchers to probe the potential range of relations that can exist between behavioural outcomes and their underlying causes, Elman et al. argue that connectionist models overturn assumptions of straightforward one-to-one mapping between mechanisms and behaviour.

Explicating knowledge Throughout their book-length investigation of what it means for something to be labelled innate, Elman et al. stress the need for researchers to be more explicit in their discussions of knowledge and knowing, and they highlight the utility of their bio-connectionist framework in achieving this end. Moreover, the characterization of knowledge that results from this exercise has important implications for developmental research. Specifically, Elman et al. argue that adopting a precise definition of knowledge that is capable of being implemented in an artificial neural network significantly undermines the case for representational nativism in development, and therefore forces researchers to rethink ways in which a behaviour could plausibly be considered 'innate'.

The role of development The authors' rejection of innate representational constraints on development has further implications for the developmental process itself. Specifically, by renouncing rich representational start states as incompatible with current neuroscientific evidence of the developing brain, Elman et al. shift the

burden of explanation for development onto the developmental process itself. They argue that an extended period of immaturity allows time for the environment to fully participate in the structuring of the maturing organism, and suggest that some complex behaviours may not be achievable without passing through a developmental process. For Elman et al., development itself is a prime causal factor in the mastery of complex cognitive abilities. As they state:

. . . development is the key to the problem of how to get complex behaviours (in the mature animal) from a minimal specification (in the genes). It's Nature's solution to the AI "scaling problem". (Elman et al., 1996, p.365)

Connectionism is not radical empiricism A commonly held view is that connectionism embodies a radical empiricist approach to human learning and development, and the authors' bio-connectionist framework comes in for similar criticism (see for example Fodor, 1997). However, as these authors are at pains to point out, all connectionist models necessarily assume some kinds of architectural and computational constraints. These constraints determine the information processing capabilities of the networks and therefore are critical to ensuring that network learning gets off the ground in each instance. Similarly, the authors' bio-connectionist perspective is not anti-nativist; their disagreement is with representational nativism more specifically. Further, they support their rejection of innate representations by undertaking an extended examination of alternative and more biologically viable sources of constraints on development, and draw directly on connectionist insights regarding the impact of differing initial architectures and timing of events to facilitate this process.

Development as emergence Finally, in addition to providing a valuable means of investigating the necessary conditions for development, connectionist models offer a vehicle for exploring the dynamics of development and in particular the issue of emergent form. Elman et al. endorse the tradition in developmental psychology and developmental biology that locates the design for final form neither in the organism

nor in the environment but in emergence from the rich interactions between them.⁵ Moreover, by advancing a dynamic framework that implements this concept in neural networks and bolsters these implementations with neuroscientific data, Elman et al. move this developmental tradition closer to a formal theory of emergent form.

4.2 Theory building in scientists and children

4.2.1 Narrowing the focus: Repositioning the child-as-scientist analogy

From the considerations highlighted above, the biologically inspired connectionism advocated by Elman et al. recommends itself as a viable theoretical alternative to the extremes of nativism and empiricism. In contrast to the theory theory, which retains a strong commitment to representational nativism, the approach of Elman et al. radically rethinks the notion of innateness and develops alternative, more biologically plausible proposals for innate predispositions, using the operation of architectural and chronotopic constraints in artificial neural networks as a guide. More generally, this bio-connectionist framework makes significant gains towards a formal theory of the processes by which innate predispositions and a structured environment interact. Whereas the theory theory is marked by the absence of any real details about the nature of these interactions, the approach of Elman et al. draws together comprehensive examples from developmental neurobiology, with concrete demonstrations of how emergent form can be implemented in connectionist models, to advance their claim that the developmental process rests on multi-level interactions. The result is a complex interactionist account of development that is broad enough to encompass the developing organism-environment relationship in its entirety. This integrated treatment further contrasts with that given by the theory theory where a reliance on theory-evidence relations to explain how change occurs restricts the theory theory's application to a narrow band of conceptual and semantic developments.

⁵ Elman et al. cite the following figures as representative of this tradition: Baldwin, Bateson, D'Arcy, Thompson, Oyama, Piaget, Vygotsky, Waddington, and Wimsatt (Elman et al., 1996, p.366).

I have suggested, then, that the interactionist perspective advocated by Elman and his colleagues offers a promising framework theory for investigations of cognitive development. Adopting this framework, moreover, has implications for the role of the child-as-scientist analogy in developmental research. Specifically, by endorsing Elman et al's interactionist approach, the debate over whether the child can be usefully conceived as an intuitive scientist takes on a decidedly narrower focus. Rather than defining the tenets of an overarching constructivist view of cognitive development as Gopnik and Meltzoff propose, the analogy is redeployed as a theoretical tool for illuminating parallels at one level of organism-environment interactions in particular. In the sections that follow, I look to reposition the child-as-scientist analogy in such a way, focusing on the specific theory building strategies employed by scientists and children in their respective attempts to achieve explanatory understanding. With this aim in mind, I turn to a body of work by Annette Karmiloff-Smith, whose detailed investigations of children's discovery processes have led her to conclude that theory building is a core component of everyday reasoning.

4.2.2 The child as a spontaneous theory builder

As one of the key contributors to the bio-connectionist framework outlined above, it is perhaps not surprising that even a cursory examination of Karmiloff-Smith's research publications reveals a strong and persistent commitment to an interactionist account of developmental change.⁶ Within this general orientation, Karmiloff-Smith has endorsed the child-as-scientist analogy, arguing that children can be usefully characterized as spontaneous theory builders:

⁶ For an early and comprehensive statement regarding the necessity of an interactionist perspective see Karmiloff-Smith (1992); for arguments highlighting the utility of integrating connectionist insights into this perspective refer to Clark and Karmiloff-Smith (1993); for recent applications of an interactionist approach to investigations of developmental disorders see Karmiloff-Smith (1998); Oliver, Johnson, Karmiloff-Smith and Pennington (2000).

For a number of years, I have argued that the child is a spontaneous theoretician, i.e. that the way in which children go about discovering how the world functions (the physical, social and linguistic worlds) is by building theories, not by simply observing facts. (Karmiloff-Smith, 1988, p.183)

Moreover, by endorsing the analogy Karmiloff-Smith looks to direct research attention away from discrete measures of behavioural performance on circumscribed tasks and towards an investigation of the processes underlying developmental change. As she states in response to a recent critique of the child-as-scientist analogy:⁷

... cognitive developmentalists originally introduced the notion of theory-in-action and explicit theory building to move beyond the mere age-related success/failure measurement of children's behaviour to try to account for the processes by which children come to achieve success. (Spencer & Karmiloff-Smith, 1997)

In keeping with this aim, Karmiloff-Smith puts the analogy to work in quite a different way from the majority of developmental researchers. Rather than exploring structural analogies between knowledge acquisition in childhood and historical development in science, Karmiloff-Smith's approach has been to advocate a process-oriented or "synchronic" perspective on the question of child-scientist parallels. According to this perspective, the focus of comparisons is on the processes underlying scientific discovery, and in particular, the changing relations between data and theory that give rise to successful problem solving.

In her 1988 article entitled 'The child is a theoretician not an inductivist', Karmiloff-Smith describes her results from a range of different problem solving situations designed to reveal these discovery processes in action. To take one example (see Karmiloff-Smith & Inhelder, 1974 for details), children were given a variety of different blocks and asked to balance them on a narrow metal support bar. The blocks differed from each other in ways that affected how they balanced on the support. That

⁷ See Gellatly (1997).

is, some blocks had their weight evenly distributed and balanced at their geometric centre. Other blocks had weights visibly attached to one end, and accordingly balanced off centre. A third type of block was “invisibly-weighted” with lead drilled into one end, so although it looked identical to the first type of block, it actually balanced off centre. By employing microgenetic analyses of children’s performance on this and other tasks,⁸ Karmiloff-Smith abstracts the following pattern of data-theory relations from children’s problem solving attempts:

- Initially, children are “data-driven”, and solve the task simply by utilizing information present in the external environment. For example, in the block-balancing task, children began by concentrating on proprioceptive information and treated each block as a new problem.
- Having succeeded at the task, children go beyond behavioural success to generate a theory that provides them with some initial explanatory understanding of the problem space. Specifically, children become “theory-driven”; they temporarily suspend their focus on solving the task and ignore or discard data that fails to fit with their nascent theory. For example, in the block-balancing task children went beyond the goal of successfully balancing the blocks to spontaneously generate a “geometric-centre theory in action” by focusing on their internal representations and rejecting data that did not agree with their theory.
- Finally, data and theory are realigned with one another and children re-engage with the original goal of successfully solving the task. However, in contrast to the pattern of initial success, this later achievement arises from the utilization of a consolidated and generalized framework. Using the block-balancing task as an example, children employed an intuitive version of the law of Torque to consider environmental feedback and successfully balance all the blocks.

⁸ Other tasks discussed include modified balance-scale problems involving the quantification of weight and length, problems investigating action and reaction as compensating forces, and spatial construction problems. See Karmiloff-Smith (1984) for more details.

Based on this robust pattern of data-theory relations that generalizes across a range of different tasks, Karmiloff-Smith concludes that children do not solve problems by progressively building up a store of atheoretical facts about the world. Rather, like scientists, their actions are constrained by powerful explanatory theories (Karmiloff-Smith, 1988).

More recently, Karmiloff-Smith has incorporated these findings into a broader conceptualization of knowledge development that details the operation of a process for increasing cognitive flexibility in human representational systems (Karmiloff-Smith, 1992, 1994). In brief, Karmiloff-Smith proposes that a fundamental aspect of human development that underscores our capacity for creative thought, involves a shift in the way in which knowledge is represented in the mind. Focusing on the status of representations that underpin different abilities and the multiplicity of levels at which knowledge is stored and available for processing, Karmiloff-Smith postulates the existence of a representational redescriptive process, whereby “. . . information that is *in* a cognitive system becomes progressively explicit knowledge *to* that system” (Karmiloff-Smith, 1994, p.694, italics in original).

According to this representational redescription hypothesis, development incorporates three recurrent phases. During Phase 1, learning is “data-driven” with the child’s achievement of “behavioural mastery” largely the result of a focus on information present in the external environment. Crucially, however, development does not end with the establishment of efficient learning procedures. Rather, successful performance is followed by an “internally driven phase”, in which the focus shifts from features of the external environment to internal representational change. As Karmiloff-Smith remarks, it is during this intermediate phase that “system-internal dynamics” predominate, and while this may result in a temporary decrement in performance, this “deterioration” is at the behavioural, as opposed to representational, level. Finally, during the third phase, internal representations and the external environmental data are “reconciled” with one another.

In terms of the format of the representations involved in this reiterative process, Karmiloff-Smith distinguishes four successive levels. At the initial level, Implicit (I), information is embedded within special-purpose procedures and is not yet available as

manipulable data to other parts of the cognitive system. In contrast, representations at the subsequent level, Explicit-1 (E1), which result from redescriptive processes operating on the implicit representations, *are* available for cross-domain applications. Specifically, Karmiloff-Smith argues that these redescribed representations “go beyond the constraints imposed at level I, where procedure-like representations are simply used in response to external stimuli” (Karmiloff-Smith, 1994, p.701). In this respect, E1 representations mark the beginnings of a cognitively flexible system, in which explicit, manipulable, and transportable representations provide the foundations for children’s early theorizing. However, while E-1 representations are available as data structures to other parts of the cognitive system, they are not yet available to consciousness. Rather, further redescription is required before knowledge becomes available to conscious access at Explicit-2 (E2) level and finally to verbal report at Explicit-3 (E3) level. On Karmiloff-Smith’s account, then, humans are active “redescribers” of their own knowledge, exploiting information that initially exists in an implicit procedural form, so that it becomes progressively explicit and accessible. It is this capacity for representational redescription that allows for flexible thought in humans, including, importantly, the ability for creative theory construction.

Concerning the issue of whether young children can be accurately characterized as theory builders, a crucial aspect of Karmiloff-Smith’s proposals is her insistence on more than two levels of representation and her rejection of the assumption that consciousness is essential for the appearance of flexible thought. As she remarks:

The most important and subtle data . . . are, in my view, those pointing to a level of representation in which knowledge is explicitly defined (i.e., represented differently from the information embedded in special-purpose domain-specific procedures of the earlier phase) but not yet available to conscious access and verbal report. Spontaneous repairs to linguistic output, unsuccessful problem solving subsequent to success, redundant behaviors, and so forth (data often ignored in developmental and adult research) are all used as vital clues to this phase of development. (Karmiloff-Smith, 1994, p.698)

By promoting a conception of E-1 level representations in which knowledge is explicitly represented and available as data to the cognitive system, but not yet available to conscious access, Karmiloff-Smith suggests a potential base for young children's spontaneous theory building. Further, Karmiloff-Smith stresses that these recurrent phases should not be thought of as age-dependent stages, signalling domain-general changes in representational format. Instead, her model posits that this representational redescription process occurs repeatedly within microdomains throughout development, and is constrained by the particular structure of the domain-specific representations on which it operates.

In summary, Karmiloff-Smith, working from an interactionist perspective on development, can be seen to fully exploit the constructivist implications of the child-as-scientist analogy as a way of investigating the processes by which children learn about the world. By granting children theories-in-action, and by focusing on the interactions between theories and data in children's problem solving, Karmiloff-Smith identifies a robust pattern of data-theory relations that supports the view that children generate theories to achieve coherent explanations of phenomena. More generally, by characterizing children's problem solving in terms of theory formation, and providing justification for this characterization via a representational redescriptive process, Karmiloff-Smith convincingly overturns the belief that theory building is the exclusive domain of meta-conceptually aware scientists. In its place, she paints a picture of theory formation as a core component of everyday reasoning, and in doing so points to a basis for mapping meaningful relations between scientists and children. As she states:

[t]he tendency to explain phenomena by a unified theory, the most general or simplest one possible, appears to be a natural aspect of the creative process both for the child and the scientist. (Karmiloff-Smith & Inhelder, 1974)

If we accept Karmiloff-Smith's claim, then the issue becomes how to best utilize the analogy with science to further our understanding of this common theory building process. And it is here that a methodological reformulation of child-scientist parallels, employing the abductive framework as a source model, offers a natural

extension of Karmiloff-Smith's proposals regarding the child as a spontaneous theoretician.

4.2.3 Detecting and explaining empirical phenomena: A framework for investigating theory building across scientific and everyday contexts

To demonstrate the potential utility of this methodological reformulation, I want to focus on the parallels between the pattern of data-theory relations uncovered by Karmiloff-Smith and the corresponding picture of these relations given by the abductive theory of scientific method. To recap briefly on the presentation given in Chapter 3, the abductive theory characterizes scientific inquiry as a problem-oriented enterprise that centres on the reliable detection and coherent explanation of empirical phenomena. According to the abductive account, inquiry typically proceeds under the direction of an evolving research problem and involves a number of interrelated phases:

- Data are initially collected and analysed to detect robust empirical regularities.
- These phenomena are then explained by abductively generating a rudimentary theory about the causal mechanisms responsible.
- Modelling strategies are subsequently deployed in an attempt to develop a more comprehensive understanding of the nature of these causal mechanisms.
- The theory is evaluated against competing alternatives on a range of criteria that emphasize the theory's capacity to provide a coherent explanatory account of the phenomena.

More generally, the abductive theory demonstrates a significant reinterpretation of data-theory relations when compared to orthodox accounts of scientific method. On the abductive method, theory construction is seen to be firmly grounded in a prior search for robust empirical regularities so that the primary inferential move is from empirical phenomena to theory. Moreover, this move is not characterized as

inductive but *abductive*, and involves going beyond a description of the phenomena to postulate causal mechanisms that are producing the phenomena. According to the abductive theory, then, scientists construct theories to explain phenomena, with empirical phenomena providing the initial stimulus for theory construction, and abductive inference assuming a pivotal role in the theory construction process.

Similarly, Karmiloff-Smith's investigations of children's problem solving strategies lead her to propose an account of spontaneous scientific discovery that endorses a data-to-theory move and focuses on the generation of explanatory understanding. More specifically, as indicated above, she identifies the following complex of recurrent interrelated phases in children's theory building attempts:

- data-focused – children display an initial concern with data, engaging in a period of rich interaction with the environment;
- theory-focused – they shift to a concern with theory, become organization oriented, and attempt to apply a single explanation to a range of disparate data;
- effective co-ordination of data and theory – data and theory are reconciled, the newly consolidated and generalized theory provides a coherent explanatory account of the problem under investigation.

Karmiloff-Smith emphasizes that this pattern of data-theory relations reoccurs repeatedly throughout development as children approach new problems in different microdomains,⁹ thereby suggesting that this domain-specific pattern is not captured by a domain-general view of scientific discovery in which children initially consider data and at a subsequent stage become capable of abstract theorizing.¹⁰

⁹ Karmiloff-Smith (1992) characterizes 'microdomains' as "subsets within a particular domain" (Karmiloff-Smith, 1992, p.6), and gives the problem of gravity in physics and the problem of pronoun acquisition in language learning as examples.

¹⁰ See also Simons & Keil (1995) for an extended argument against a domain-general shift in development from concrete to abstract thought.

Moreover, Karmiloff-Smith argues that this data-to-theory move has an *explanatory* function in children's problem solving.¹¹ In fact, the orienting focus of her 1988 paper is to substantiate the claim that the child is a spontaneous theoretician by contrasting both children's and scientists' theory building activities with a naïve inductivist account of scientific inquiry. On this account, as we saw in Chapter 2, scientific inquiry proceeds by gathering all the facts in a non-judgmental manner to retain their objectivity and then applying the rules of inductive inference to arrive at generalizations or descriptive laws about those facts. Karmiloff-Smith's investigations however, lead her to reject inductivism as an accurate picture of children's scientific practice:

Both for the child and the adult researcher, scientific progress does not stem from the use of logical criteria on the basis of rational induction from observations. (Karmiloff-Smith, 1988, p.183)

Furthermore, having generated their initial explanatory theories, Karmiloff-Smith argues that children do not attempt to falsify them, as Popper's hypothetico-deductive account of scientific inquiry would dictate. Instead, she repeatedly emphasizes the tenacity of children's theories and their constraining influence on children's problem solving behaviour on the tasks, describing how children will ignore "glaring counter examples" to their theories and even invent 'observable' data in order to maintain their theoretical commitments. Based on such findings, then, it would seem that the logic-and-testing model at the heart of hypothetico-deductive method also fails to capture the process of spontaneous scientific discovery in children:

... clearly children are not falsificationists. They constantly develop theories and create domains, carving and re-carving nature at new joints. And they simplify and unify incoming data to make them conform to their theories. (Karmiloff-Smith, 1988, p.192)

¹¹ See also Keil (1998) for the speculation that explanation may be more central to cognitive development than prediction.

Based on the above analysis, it seems clear Karmiloff-Smith endorses the view that the component processes of scientific discovery are not qualitatively distinct from those seen in children's problem solving. However, when she looks to an account of scientific inquiry to illuminate these parallels, she finds that neither inductive nor hypothetico-deductive accounts are adequate source models. From the standpoint of abductive method however, the pattern of data-theory relations in children's problem solving that she describes can be seen to bear some striking parallels to the picture of these relations defended by Haig (2002). Firstly, Karmiloff-Smith's account of children's actual scientific practice alludes to the activity of phenomena detection by indicating that children initially exploit information present in the environment to identify empirical regularities or patterns in the data. Secondly, the generative move from data (phenomena) to theory in children's problem solving is best described as abductive in character. Instead of simply observing facts and reasoning via inductive logic to descriptive generalizations about them, Karmiloff-Smith suggests children spontaneously generate rudimentary explanatory theories to make sense of the data patterns they are confronted with, by going beyond the data to identify causes that explain these patterns or 'effects'.

Further, concerning the strategies available to children for developing their theoretical explanations, Karmiloff-Smith highlights the role of various modelling strategies in children's theory building. For example, in a spatial construction task, children were observed explicitly employing symmetry as a heuristic to potentiate their search. Likewise, children were also seen to use quantification as a promising search path, counting elements in a physics problem to see if the resulting pattern was suggestive of the causal entities at work (Karmiloff-Smith, 1984). Finally, while not dealing explicitly with the question of how (or if) children evaluate their theories, Karmiloff-Smith explicitly rejects the idea that children reason like falsificationists. Moreover, in emphasizing how children go beyond behavioural success to spontaneously restructure their knowledge base, she suggests that it is the achievement of a coherent explanatory framework that signals to children that they have successfully solved the problem under study.

4.3 Beyond the theory theory

In this chapter, I have attempted to establish the initial plausibility of the proposal that the abductive framework can inform deliberations about the knowledge construction strategies of young children. In order to achieve this end, I began by considering the intended scope of Gopnik and Meltzoff's theory theory to the development of knowledge in childhood. I argued that these researchers employ the analogy with science as a means of overcoming the limitations of existing theoretical frameworks, and look to a particular model of scientific change to define a viable alternative theory of cognitive development. However, given the inability of the theory theory to offer any substantive ideas about the processes through which children's theories are developed, I suggested that this formulation fails to cultivate meaningful correspondences between children and scientists, and in doing so indicates the need to consider alternative frameworks for developmental inquiry.

Following this recognition, I turned to examine a broad interactionist perspective on development advanced by Elman and colleagues. I demonstrated how these researchers draw on recent advances in the neurosciences and in computational modelling to define a bio-connectionist approach, that is neither radically nativist nor radically empiricist in its commitments. Specifically, this integration of biology and connectionism affords a unique view of development as a cascade of multi-level interactions, and is therefore capable of encompassing all aspects of the organism-environment relationship that are relevant to an understanding of developmental change. I concluded that the work of Elman et al. offers researchers a promising theoretical framework for investigations of cognitive development, and a superior alternative to the theory theory model, which demonstrates only limited application to a restricted set of developmental issues.

Having endorsed this interactionist framework, I subsequently set about repositioning the child-as-scientist analogy in order to focus on the specific theory building strategies of children and scientists. Drawing on the detailed investigations of children's problem solving undertaken by Karmiloff-Smith, I highlighted some significant parallels between the pattern of data-theory relations identified by Karmiloff-Smith, and the corresponding picture of these relations articulated by

abductive method. Based on these parallels, I believe it is plausible to argue that this methodological reformulation has the potential to move the analogy beyond the current theory theory model and develop system mappings of higher order relations between children and scientists. With this aim in mind, therefore, Chapter 5 extends the abductive-methods analogy to the context of theory evaluation and examines the processes by which children evaluate the quality of their explanations.

Chapter 5

Evaluating Ideas: The Role of Coherence in Scientific and Everyday Thought

In Chapter 4, I sought to establish the initial plausibility of my methodological reformulation of the child-as-scientist analogy by identifying an abductive pattern of reasoning in accounts of children's spontaneous theory building. Given evidence that children generate explanatory theories, and do so by reasoning abductively, the focus of this chapter turns to theory evaluation; that is, how children appraise the quality of everyday explanations, and whether the cognitive strategies they employ bear any significant resemblance to those employed by scientists.

Accordingly, in Section 5.1 I begin by reviewing recent comparisons of theory use in scientific and everyday contexts that offer both indirect and direct evidence for a substantial overlap in the criteria used by scientists and children to evaluate explanations. Section 5.2 ties these findings to the computational model of theory evaluation advanced by Paul Thagard, which demonstrates how these criteria are combined in inferences to the best explanation on the basis of explanatory coherence. In Section 5.3, I conclude that a methods-centred approach to theory evaluation that incorporates precise computational models of the mechanisms underlying explanatory reasoning, reveals significant relations between the cognitive strategies employed in scientific and everyday thought.

5.1 Employing theories in scientific and everyday contexts

In the previous chapter, I repositioned the child-as-scientist analogy to focus on the specific methods and strategies employed by children in their knowledge-seeking endeavours. Using the abductive theory of scientific method as a source model, I argued for meaningful relations between the particular sequence of data-theory relations underpinning children's theory generation efforts, and the codified account of this generative process provided by the abductive theory. Given these suggestive parallels between children and scientists in the context of theory generation, this chapter further develops the claim that the abductive framework offers a source of ideas about children's knowledge construction strategies, by investigating the role of abductive reasoning in the context of evaluation.¹

Concerning the issue of how children evaluate ideas, is it plausible to suggest they reason like scientists? That is, when faced with competing explanations of a phenomenon, do children possess the cognitive skills to make a rational judgment about which alternative provides the best explanation of the phenomenon in question? And if so, are their judgments constrained by the same sorts of criteria that feature in contemporary models of scientific theory evaluation?

While such questions would seem central to comparisons of theory use in scientific and everyday contexts, there is little research to date that bears directly on these issues. With the exception of the process-oriented investigations conducted by Karmiloff-Smith reviewed in Chapter 4, the majority of studies examining children's commonsense theories have focused on the underlying theoretical frameworks that are argued to constrain core knowledge in foundational domains (e.g., Carey, 1985; Keil, 1989; Wellman, 1990; Wellman & Gelman, 1992, 1998). While this emphasis on

¹ I have chosen to focus specifically on theory evaluation rather than theory development issues, because theory theory analyses have tended to concentrate efforts on theory formation and revision, as opposed to theory development. As I indicated in Chapter 4, indirect evidence from Karmiloff-Smith's studies demonstrating children's effective use of heuristics on problem solving tasks is suggestive of the early use of modelling strategies in theory construction. However, see Wilson and Keil (2000) for the proposal that everyday explanations employed by children and lay adults demonstrate a lack of systematic development when compared to scientific explanations.

framework theories has been helpful for determining the ways in which knowledge is structured and organized within domain-specific systems, questions about the specific strategies children use to evaluate the adequacy of their knowledge, and the similarities and/or differences to strategies employed by scientists, have been relatively neglected. One notable exception is a body of research investigating children's theories of the natural world undertaken by William Brewer and colleagues (Brewer & Samarapungavan, 1991; Brewer, Chinn, & Samarapungavan, 2000; Samarapungavan, 1992; Samarapungavan & Wiers, 1997; Vosniadou & Brewer, 1987, 1992, 1994), which provides some informative comparisons of theory use in scientific and everyday contexts. In what follows, I draw on both indirect and direct evidence from this body of work to examine the strategies children use to choose between competing explanations of phenomena, and whether these evaluative strategies bear any significant relation to theory evaluation practices in science.

5.1.1 Children's theories of the natural world: Examples from observational astronomy

Undertaking an analysis of possible child-scientist parallels in the context of theory evaluation, presupposes not only that children construct theories about the world on the basis of their everyday experience, but also that these knowledge structures display, at least in part, what we take to be crucial features of theory goodness. In Chapter 3, I highlighted a range of criteria that are generally taken to be important for evaluating the worth of scientific theories, including empirical accuracy, consilience or explanatory breadth, simplicity, and consistency (Haig, 2002; Thagard, 1992). An initial question, then, is whether children's theories also display these features.

Coherence versus fragmentation

A review of descriptions of children's knowledge structures by researchers who endorse some version of the analogy with science (e.g., Carey, 1985; Karmiloff-Smith, 1988; Karmiloff-Smith & Inhelder, 1974; Wellman, 1990; Gopnik & Wellman, 1994; Gopnik & Meltzoff, 1997), reveals an emphasis on the *coherence* and *systematicity* of children's intuitions. Moreover, these features are often explicitly employed as diagnostic criteria for determining the theoretical status of children's

knowledge (Wellman, 1990; Gopnik, 1996b; Gopnik & Meltzoff, 1997). In addition, researchers have also stressed the *consilience* of children's theoretical constructions and the importance of the resulting explanatory understanding for successful problem solving (Karmiloff-Smith, 1988; Keil, 1998). Similarly, science education researchers who propose that intuitive knowledge has the status of a naïve theory, have highlighted the *explanatory power* of these alternative frameworks to integrate and make sense of a wide range of disparate phenomena (e.g., Driver & Easley, 1978; Kempton, 1987; McCloskey & Kargon, 1988). Further, they argue that this feature contributes to the well-documented resistance of intuitive ideas to formal science instruction.

In contrast, other researchers who reject the idea that initial knowledge is theory-like (e.g., diSessa, 1983, 1988, 1993; Solomon, 1983, 1996), have tended to support their claims by highlighting the absence of these features in children's cognitive constructions. For example, diSessa (1988) has argued that intuitive knowledge is typically fragmented and inconsistent, and does not demonstrate the explanatory breadth or depth characteristic of scientific theories. Taking naïve physics as a case study, he proposes that everyday intuitions about the physical world are best described as loose assortments of discrete pieces of information or knowledge fragments, rather than coherent explanatory theories. According to diSessa, these intuitions are simple abstractions from everyday experience and tend to be inconsistently applied in problem solving (diSessa, 1988).

In light of such fundamental disagreements about the character of intuitive knowledge, a large-scale research project investigating children's theories of astronomical phenomena (Vosniadou & Brewer, 1992, 1994) is informative. Working from the assumption that children are spontaneous theory builders, these researchers have undertaken detailed examinations of children's mental representations of various phenomena that form part of the domain of scientific astronomy. Specifically, Vosniadou and Brewer (1992, 1994) conducted structured interviews in which elementary school children were asked a series of factual and generative questions about the shape of the earth and the day/night cycle. Based on children's verbal responses and their associated drawings, Vosniadou and Brewer attempted to identify

the mental models held by elementary school children, and to determine whether these models were applied in a consistent manner across a range of problems.

The results of these studies indicate that even young children develop coherent explanatory theories of natural world phenomena, which they apply in a systematic and consistent fashion. In the case of children's models of the earth, Vosniadou and Brewer (1992) identified five well-defined alternative models to the spherical earth model, depicted in Figure 5.1.

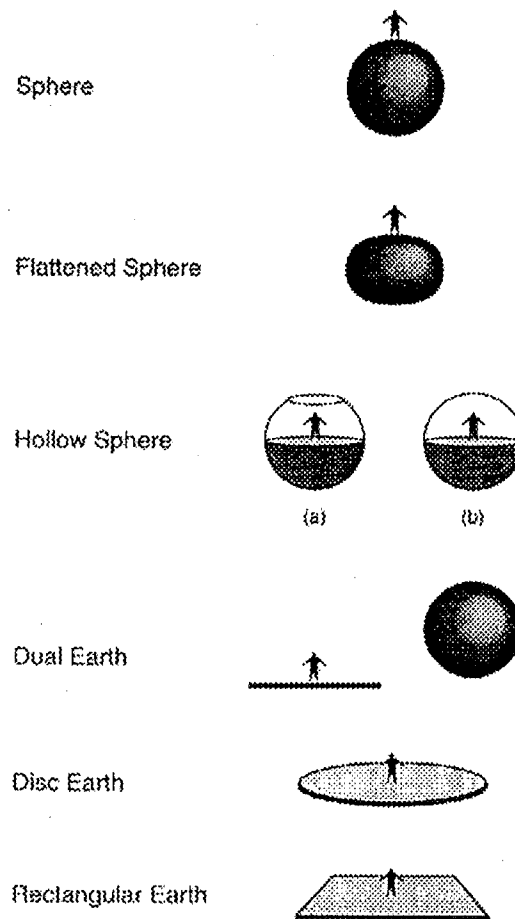


Figure 5.1 Children's mental models of the earth. Reproduced from Vosniadou and Brewer (1992, p.549).

The responses of the youngest children (first graders) suggested that many of them held flat earth models (i.e., disc earth and rectangular earth). These models, in which the earth is supported by ground, were in keeping with children's everyday experience and were unaffected by the culturally accepted spherical earth model. At the other end of the spectrum, many of the older children's (fifth graders) protocols indicated that they had formed the culturally accepted scientific model in which the earth is a sphere surrounded by space. In addition to these models, Vosniadou and Brewer (1992) also uncovered a number of creative "synthetic models" of the earth (dual earth, hollow sphere, flattened sphere), which were formed by a significant proportion of the participants.

Close examination of these models by the researchers revealed they were the outcome of a concerted attempt by children to reconcile their everyday experience of a flat earth with culturally conveyed information that the earth is a sphere (Vosniadou & Brewer, 1992). For example, children who constructed the dual earth model demonstrated the belief that there are two earths: one that is flat, where people live, and another one that is round and located up in the sky with other planets. Evidence that children held this dual earth model and were not merely giving confused or inconsistent responses to questions came from children's drawings in which both 'earths' were represented, and from the use of two distinct terms (i.e., the term 'earth' was typically reserved for the round planet up in the sky, while the term 'ground' was used to refer to the place where people live) (Vosniadou & Brewer, 1992).

In addition to the dual earth model, a significant proportion of children were classified as holding a hollow sphere model. These children indicated that they had resolved the flat earth/sphere conflict by forming a well-defined model of the earth as either a) a hollow sphere with people living on flat ground deep inside it, or b) as consisting of two hemispheres, with people living on the lower hemisphere and the sky forming the upper hemisphere in the shape of a dome. Vosniadou and Brewer (1992) offer the following protocol and associated drawing as representative of children who were found to have constructed this creative synthetic model:

Venica, Grade 3 (hollow sphere):

- E: How come here the earth is flat but before you made it round?
 C: Because you are on the ground and you make that picture like a shape and you made it a square shape and if you'll look up it'll look like a rectangle or something like that and if you go out of earth and go into space you'll see a circle or round.
 E: So what is the real shape of the earth?
 C: Round.
 E: Why does it look flat?
 C: Because you are inside the earth.
 E: If you walked and walked for many days in a straight line, where would you end up?
 C: Somewhere in the desert.
 E: What if you kept walking?
 C: You can go to states and cities.
 E: What if you kept on walking?
 C: (No response.)
 E: Would you ever reach the edge of the earth?
 C: No. You would have to be in a spaceship if you're going to go to the end of the earth.
 E: Is there an edge to the earth?
 C: No. Only if you go up.
 Later:
 E: Can people fall off the end/edge of the earth?
 C: No.
 E: Why wouldn't they fall off?
 C: Because they are inside the earth.
 E: What do you mean inside?
 C: They don't fall, they have sidewalks, things down like on the bottom.
 E: Is the earth round like a ball or round like a thick pancake?
 C: Round like a ball.
 E: When you say that they live inside the earth, do you mean they live inside the ball?
 C: Inside the ball. In the middle of it.

(Vosniadou & Brewer, 1992, pp.563-4)

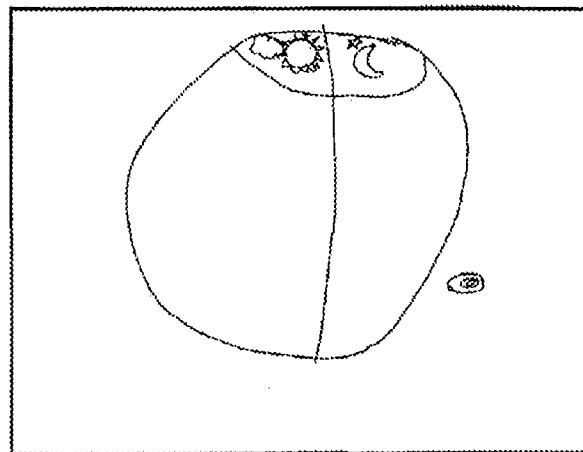


Figure 5.2 Drawing of the earth, moon, the stars, and the sky by Venica, Grade 3 (hollow sphere). Reproduced from Vosniadou and Brewer (1992, p.558).

Finally, children who were classified as holding a flattened sphere model of the earth were found to have integrated information that the earth is round with their experience of its flatness, by forming a model of the earth as a thick pancake. According to this model, the earth is rounded at the sides and flat on the top and bottom where people live, supported by gravity (Vosniadou & Brewer, 1992).

The above sample of findings, resulting from Vosniadou and Brewer's detailed analyses of children's mental models of the earth, can be taken to support the view that children construct coherent explanatory theories of natural world phenomena. Contrary to diSessa's claim that children's intuitive knowledge is fragmented and unconnected, Vosniadou and Brewer (1992) found that the great majority of children interviewed had recourse to a well-defined model of the earth, which they applied consistently. Specifically, by assuming that the children in their study were using a small number of clearly defined models of the earth, Vosniadou and Brewer were able to account for over 80% of the variation in children's individual responses. In addition, Vosniadou and Brewer point to the high frequency of synthetic models uncovered as a powerful indicator that children are not content with fragmentary knowledge about the earth's shape but instead will take active steps to integrate the information they receive into a coherent representation (Vosniadou & Brewer, 1992). Similar findings of coherence in studies of children's models of the day/night cycle (Vosniadou & Brewer, 1994), and their ideas about speciation (Samarapungavan & Wiers, 1997), add further weight to Vosniadou & Brewer's claims for theory construction in childhood.

More generally, Vosniadou and Brewer's findings that a large number of children formed *alternative* models of the earth, offers a potential explanation for the perceived fragmentation and inconsistencies in children's intuitive knowledge discussed by researchers such as diSessa (1988). As Vosniadou and Brewer (1992) make clear, some of the models uncovered, such as the hollow sphere model, were so novel from the standpoint of the culturally accepted scientific model, that children initially appeared to be confused and inconsistent in their responses. Vosniadou and Brewer caution that without detailed characterizations of the actual models employed in a particular domain, researchers may mistakenly attribute inconsistencies and self-

contradictions to children that are unwarranted when their conceptual frameworks are fully appreciated and taken into account (Vosniadou & Brewer, 1992).

Assessing children's theories: Criteria of theory goodness

If it can be accepted that children construct coherent explanatory theories of natural world phenomena, it is plausible to consider whether children's knowledge structures also meet criteria employed to evaluate the adequacy of theories in science. As part of their investigation of children's models of the day/night cycle, Vosniadou and Brewer (1994) explicitly examined three such criteria: *empirical accuracy*, *logical consistency (both internal and external)*, and *simplicity*. Firstly, concerning the criterion of empirical accuracy, they found good evidence to support the claim that children honour the need for empirically consistent models in constructing their representations of the day/night cycle. Specifically, Vosniadou and Brewer argued that the majority of children's models were consistent with the following range of observations available to young children:

- There is a sequence of day and night.
- The sun is in the sky during the day, but not at night.
- The moon and stars are in the sky during the night but not during the day.
- Objects appear and disappear.

(Vosniadou & Brewer, 1994, p.130)

The second criterion used by Vosniadou and Brewer (1994) to evaluate children's day/night cycle models, focused on the degree to which these models were internally and externally consistent. Once again, the detailed representations of children's models built up from verbal protocols and elicited drawings, offer considerable evidence that even young children are sensitive to the criterion of logical consistency. In particular, Vosniadou and Brewer found that the majority of the children interviewed gave consistent responses to questions about a range of different phenomena associated with the day/night cycle such as the disappearance of the sun at night and the disappearance of the stars during the day (Vosniadou & Brewer, 1994).

Moreover, by investigating children's models of the earth together with their models of the day/night cycle and their models of the sun, these researchers were able to determine the degree to which children's accounts of related phenomena demonstrated external consistency with one another. Specifically, their findings indicated that for the majority of children interviewed, their models of the earth and the sun constrain the particular models they develop of the day/night cycle, and do so in appropriate ways. For example, children who held a flat earth model in which the earth is rooted in the ground, did not go on to explain the day/night cycle by invoking the mechanism of a mobile earth rotating on its axis or revolving around the sun. Instead, they opted for a logically consistent mechanism such as a movable sun that goes behind a mountain or far out into space at night and returns during the day. Similarly, children who were found to hold a stationary sun model did not subsequently explain the day/night cycle in terms of the movement of the sun. Rather, these children relied on logically consistent mechanisms such as the occlusion of the sun by clouds or the earth revolving around a stationary sun (Vosniadou & Brewer, 1994).

Finally, in addition to empirical accuracy and logical consistency criteria, Vosniadou and Brewer (1994) also examined the extent to which children demonstrate an awareness of *simplicity* in their explanations of astronomical phenomena.² They found that the majority of children not only employed mechanisms that were logically consistent, but in many cases used a single mechanism to explain multiple phenomena. For example, many children were shown to use the same mechanism to account for both the disappearance of the sun and the disappearance of the moon in their protocols. More generally, Vosniadou and Brewer (1994) argue that children's models, such as that depicted in Figure 5.3, are often much simpler than the accepted scientific model, supporting the idea that children employ a simplicity measure in their theory construction efforts.

² In the context of their study, Vosniadou and Brewer defined 'simplicity' as "... the use of the same mechanism to account for different, although related phenomena" (Vosniadou & Brewer, 1994, p.176).

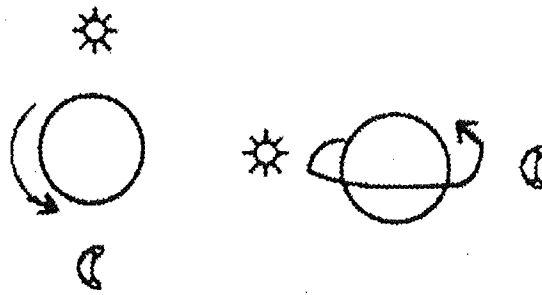


Figure 5.3 Example of a mental model of the day/night cycle constructed by children in Vosniadou and Brewer's study that is simpler than the scientific model. In this model, the sun and moon are fixed at opposite sides of the earth, which rotates in an up/down or east/west direction. Reproduced from Vosniadou and Brewer (1994, p.169).

Collectively, the investigations of children's mental models of astronomical phenomena conducted by Vosniadou and Brewer (1992, 1994) offer indirect support for the possibility of parallels between children and scientists in the context of theory evaluation. By demonstrating that children's intuitive knowledge displays many of the features deemed to be hallmarks of good theories in science, this research suggests that children are capable of constructing and reasoning from knowledge structures that conform to these criteria.³ A further study conducted by Ala Samarapungavan (Samarapungavan, 1992), which does give explicit attention to theory evaluation, suggests that children can also use these criteria to evaluate the adequacy of their knowledge. The following section reviews Samarapungavan's important results.

³ This claim is made more compelling by the nature of the domain investigated. That is, while children's explanatory models of the earth and the day/night cycle are likely to be constrained by some core physical principles, it is not plausible to suggest that the particular models uncovered by Vosniadou and Brewer, with their demonstrated empirical adequacy, logical consistency, and simplicity characteristics, are part of our innate endowment (as is often argued in the case of folk psychology for example). Vosniadou and Brewer make a similar point regarding the support their findings lend to a strong constructivist position on knowledge development more generally (Vosniadou & Brewer, 1994).

5.1.2 Evaluating explanations: Criteria for theory choice

In an attempt to determine whether it is useful to cast children as intuitive scientists, Samarapungavan (1992) investigated the reasoning strategies employed by elementary school children on a range of theory choice tasks. In particular, she was concerned to establish whether children have access to and can utilize the same kinds of metaconceptual criteria that underpin theory evaluation in science. Working from accounts of theory selection in contemporary philosophy of science (e.g., Kuhn, 1977; Laudan, 1977; Popper, 1959; Thagard, 1978; Toulmin, 1972), Samarapungavan identified the following four criteria for further investigation with children:

- Range of explanation: Firstly, if children reason like scientists then they should demonstrate a preference for theories that can account for a greater range of the data to be explained.
- Non-ad hocness: Secondly, children should prefer theories that explain the data without having to rely on additional auxiliary or ad hoc hypotheses.
- Empirical consistency: Thirdly, if children's judgments mirror scientists' judgments on theory choice tasks, then they should select theories that are consistent with the empirical evidence over theories that are inconsistent with it.
- Logical consistency: Finally, Samarapungavan proposes that children should also demonstrate a sensitivity to the logical consistency of theories. That is, if a theory is shown to have internal inconsistencies then this fact should enter into, and constrain, children's choices on theory selection tasks (Samarapungavan, 1992).

Having identified these four criteria as pivotal to scientific theory evaluation, Samarapungavan (1992) devised a series of experimental tasks to test children's ability to use these criteria as a basis for theory selection. In each task, she presented elementary school children with a group of observations relating to a particular

phenomenon. The children were then given two alternative explanations for the observations that differed on the basis of one of the four criteria listed above and were asked to choose the 'correct' explanation from the two alternatives. Children were also asked to justify their theory choices. Finally, in order to examine the influence of children's prior beliefs on their ability to apply the criteria under examination, Samarapungavan systematically manipulated the content of the tasks so that they were consistent, inconsistent, or neutral with respect to children's prior knowledge. Therefore, each of the four criteria was tested with three sets of materials: two sets from the domain of astronomy where pre-testing determined children's adherence to either a geocentric or heliocentric framework, and one set from the domain of chemistry which was conceived to be knowledge neutral (Samarapungavan, 1992).

To give an example of the experimental materials used, consider the following set-up that was used to test children's ability to apply the 'range of explanation' criterion and the 'non-ad hocness' criterion (Figure 5.4).

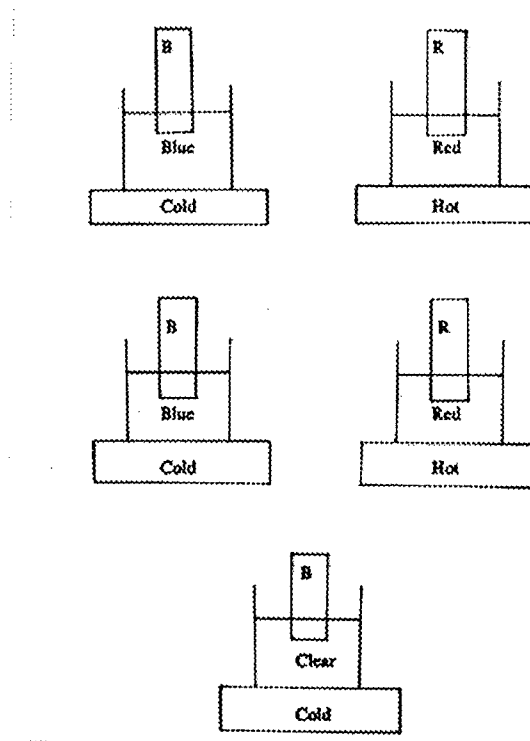


Figure 5.4 Experimental materials used to test children's application of the range of explanation criterion and the non-ad hocness criterion. Reproduced from Samarapungavan (1992, p.12).

In this chemistry task, a collection of jars was mounted on boxes that were labelled either 'hot' or 'cold'. Each jar was filled with a blue, red, or colourless liquid (acids and bases). A pH indicator was immersed into each jar, and its change in colour noted. In Figure 5.4, 'B' represents a change to the colour blue, while 'R' represents a change to the colour red (Samarapungavan, 1992).

In tests of children's use of the 'range of explanation' criterion, children observed these changes and then were given two alternative theories (T1 and T2), that supposedly corresponded to what two children their own age (Ann and Joe) thought about the observations:

T1 – I think the stuff in the jars is paint. The stick is coated with the colour of the liquid. So when you put the stick in the blue paint it turns blue and when you put the stick in the red paint it is painted red.

T2 – I think that the stick changes color to show if a thing is hot or cold. So when the liquid in the jar is heated by the hot box, the stick turns red. The stick even turned blue in the liquid that had no colour because the jar was cold. (Samarapungavan, 1992, p.11)

T1 attempted to explain the observations by proposing that the liquids were dyes, but was unable to account for fact that the indicator stick in the colourless liquid turned blue. In contrast, T2 was able to account for all the observations by proposing that the stick 'measured' the temperature of the liquids, turning blue in cold liquids and red in hot ones (Samarapungavan, 1992).

In tests of children's use of the 'non-ad hocness' criterion, T1 introduced an ad hoc explanation in an attempt to account for the outstanding observation that the indicator stick also changed colour in the clear liquid:

T1 – I still think the stick changes colour because it is covered by the paint in the jar. So when you put it in the red paint it turns red and in the blue paint it turns blue. Only sometimes, if the sticks are old they get

spoiled and they begin to get blue spots like the one in the jar with the clear liquid. (Samarapungavan, 1992, p.14)

The results of Samarapungavan's (1992) study offer convincing evidence that even young children can use a range of metaconceptual criteria to evaluate competing accounts of phenomena. Specifically, Samarapungavan found that elementary school children could use all four criteria investigated as a basis for theory selection when the theories in question did not conflict with their existing knowledge.⁴ Even the youngest children interviewed (first graders) demonstrated a systematic preference for theories that explained more, and were consistent both internally and with the evidence. In addition, the theory choices of older children (fifth graders) also showed a systematic preference for theories that did not rely on ad hoc hypotheses to account for the observations presented. Further, in many cases children were able to justify their choices by explicit reference to the criterion being examined.⁵ To give an example of a criterion-based justification on the 'range of explanation' task described above, Samarapungavan recounts the following justification given by a child after her selection of T2 (the theory of broader range):

... I think Ann is right because she also showed why the stick is blue in that box [points to clear jar] and Joe didn't. (Samarapungavan, 1992, p.13)

Based on these findings, Samarapungavan argues that children can evaluate ideas by applying a range of metaconceptual criteria, and do so in ways that resemble theory evaluation strategies employed by scientists. She concludes that contrary to claims made by critics of the child-as-scientist analogy (e.g., Kuhn, 1989), "... even young children share some of the cognitive underpinnings of scientific rationality that scientists do" (Samarapungavan, 1992, p.1).

⁴ Overall, the frequency of correct choices was very high: pooled across grade, 1267 out of a total 1620 theory choices or 78.2% were for the correct theory. For each criterion, the percentage of correct choices pooled across grade and task were as follows: range of explanation (81%), non-ad hocness (65%), empirical consistency (94%), logical consistency (82%) (Samarapungavan, 1992).

⁵ This was particularly true of the 'range of explanation' and 'empirical consistency' conditions, where criterion-based justifications for correct theory choices were 86% and 96% respectively (Samarapungavan, 1992).

To summarize, Samarapungavan's (1992) findings of scientific rationality on theory choice tasks, when taken together with the results of Vosniadou and Brewer's (1992, 1994) investigations of children's theories of astronomy, indicate the possibility of significant relations between children and scientists in the context of theory evaluation. Indirect evidence from studies of children's mental models of the earth and the day/night cycle, suggests that children construct coherent explanatory theories that conform closely to criteria used to assess the worth of scientific theories. Direct evidence that children can actually use these criteria to choose between competing theories, and often justify their choices by reference to these criteria, adds considerable weight to the claim for parallels between the strategies employed by children and scientists to evaluate the efficacy of ideas.

In a recent review of the nature of explanation in children and scientists (Brewer, Chinn & Samarapungavan, 2000), Brewer and colleagues draw together these findings concerning children's theories of the natural world to provide an account of explanation and the criteria children use to evaluate the quality of explanations. Based on their research, they conclude that there is a considerable overlap in the evaluative criteria used by both groups to assess the adequacy of knowledge, and speculate on how to further understanding of these correspondences in theory application across scientific and everyday contexts. Based on the methodological reformulation of child-scientist parallels advocated in this thesis, it is suggested that these findings are naturally interpreted within the computational model of theory evaluation endorsed by the abductive theory of scientific method, according to which multiple criteria like those identified by Brewer et al. are combined in inferences to the best explanation on the basis of explanatory coherence.

5.2 Theory building, evaluation, and explanatory coherence

In Chapter 3, as part of my presentation of the abductive theory of scientific method, I highlighted Haig's commitment to a multi-criterial perspective on theory evaluation, and more specifically to the process of evaluating theories in terms of their explanatory power or worth (Haig, 2002). Adopting Gilbert Harman's label of

'inference to the best explanation' (Harman, 1965), Haig argues that theory evaluation in science is best construed as a comparative exercise, whereby rival theories are pitted against each other on multiple dimensions to determine which theory provides the best explanation of the empirical phenomena under study. Seeking a more detailed account of how this process might operate in scientific practice, Haig (2002) discusses an influential formulation of inference to the best explanation developed by the cognitive scientist Paul Thagard (Thagard, 1989, 1992) that centres on the notion of *explanatory coherence*.

5.2.1 The theory of explanatory coherence (TEC)

According to Thagard, the evaluation of competing theories or hypotheses is decided on the basis of three main criteria that collectively determine their relative explanatory coherence. The first factor involved in assessments of explanatory coherence concerns how much of the evidence a particular hypothesis explains, a criterion that Thagard terms the *explanatory breadth* or *consilience* of a hypothesis (see also Whewell, 1967). The second factor important to explanatory coherence considerations is whether its explanations are economical and free of ad hoc assumptions. Thagard takes this factor to be concerned with issues of *simplicity*. The third factor identified by Thagard as contributing to explanatory coherence is the degree to which the hypothesis in question is similar to hypotheses that explain similar phenomena, a criterion that Thagard labels *analogy*.⁶ Thagard argues that when these three criteria are combined they serve to determine the explanatory coherence of a hypothesis relative to available alternatives. That is, a theory is judged to be more explanatorily coherent and therefore provides a better explanation of the phenomena than its rivals if it explains more, relies on fewer ad hoc hypotheses to achieve this explanatory success, and is consistent with currently accepted theories that explain similar phenomena (Thagard, 1992). In developing in detail his theory of explanatory coherence and its implementation in a computer program, which he applies to both

⁶ An additional consideration is whether the hypothesis itself is explained, which Thagard subsumes under the criterion of explanatory breadth (Thagard, 1992).

scientific and everyday reasoning, Thagard endeavours to show how such judgments are made.

With this aim in mind, Thagard (1992) compiles the following list of seven principles. Taken together, these principles serve to establish relations of explanatory coherence and allow a judgment of the acceptability of propositions that comprise an explanatory system, as well as an assessment of the explanatory coherence of the system as a whole, on the basis of local relations holding between pairs of propositions. In the following presentation of the principles, the symbol 'S' stands for an explanatory system, which consists of propositions P, Q, and $P_1 \dots P_m$ (refer Thagard, 1992, pp.65-69; for an informal statement of the principles see Thagard, 2000, p.43):

1. Symmetry

- (a) *If P and Q cohere, then Q and P cohere.*
- (b) *If P and Q incohere, then Q and P incohere.*

The first principle, Symmetry, simply states that explanatory coherence between two propositions is a symmetrical relation.

2. Explanation

If $P_1 \dots P_m$ explain Q, then:

- (a) *For each P_i in $P_1 \dots P_m$, P_i and Q cohere.*
- (b) *For each P_i and P_j in $P_1 \dots P_m$, P_i and P_j cohere.*
- (c) *In (a) and (b) the degree of coherence is inversely proportional to the number of propositions $P_1 \dots P_m$.*

Principle 2, Explanation, determines the majority of the explanatory relations that give rise to explanatory coherence and in doing so, subsumes the criteria of explanatory breadth and simplicity discussed above. In particular, it specifies that (a) if a hypothesis explains a piece of evidence (or another hypothesis) then it coheres with that evidence (or hypothesis); (b) if two hypotheses jointly explain something i.e., they are "co-hypotheses", then they cohere with each other; and (c) the greater the number of hypotheses required for an explanation, the lower the degree of coherence of the hypotheses with each other and with what is being explained.

3. Analogy

If P_1 explains Q_1 , P_2 explains Q_2 , P_1 is analogous to P_2 , and Q_1 is analogous to Q_2 , then P_1 and P_2 cohere, and Q_1 and Q_2 cohere.

The third Principle, Analogy, embodies the other major criterion identified by Thagard as central to establishing explanatory coherence. This principle states that if similar propositions explain similar pieces of evidence, then they cohere with each other.

4. Data Priority

Propositions that describe the results of observation have a degree of acceptability on their own.

This principle recognizes that the results of observation and experiment, while not indubitable, have some independent acceptability. In science, this independent warrant can be seen to derive from the application of data collection and analysis techniques that are designed to promote data reliability (Haig, 2002). By including this principle as part of his theory of explanatory coherence, Thagard therefore endorses a form of “discriminating coherentism” (Thagard, 2000, p.44), where data is given a certain priority in assessments of explanatory coherence, but can be overridden by coherence considerations if required.

5. Contradiction

If P contradicts Q , then P and Q incohere.

Principle 5, Contradiction, covers the negative relations that hold between contradictory hypotheses and states that if two propositions contradict each other, then they will “incohere” or actively resist cohering.

6. Competition

If P and Q both explain a proposition P_i , and if P and Q are not explanatorily connected, then P and Q incohere. Here P and Q are explanatorily connected if any of the following conditions holds:

- (a) P is part of the explanation of Q ,*
- (b) Q is part of the explanation of P ,*
- (c) P and Q are together part of the explanation of some proposition P_j .*

While Principle 5 relates specifically to contradictory hypotheses, Principle 6 covers all other cases where hypotheses are deemed incompatible because they compete to explain the same evidence and no explanatory relations hold between them.

7. *Acceptability*

- (a) *The acceptability of a proposition P in a system S depends on its coherence with the propositions in S .*
- (b) *If many results of relevant experimental observations are unexplained, then the acceptability of a proposition P that explains only a few of them is reduced.*

The final principle included in Thagard's (1992) statement of TEC proposes that propositions are accepted or rejected based on their degree of coherence with other propositions, as established by Principles 1-6. In doing so, this principle embodies the fundamental assumption driving assessments of explanatory coherence, namely that the decision to accept or reject a theory as a whole is made on the basis of local pair-wise coherence relations (Thagard, 1992).

5.2.2 Computing explanatory coherence: An introduction to ECHO

While the principles of explanatory coherence listed above offer a far more precise formulation of the notion of inference to the best explanation than was previously available, Thagard argues that the theory by itself is still too general to show exactly how it is possible to compute the acceptability of competing hypotheses or theories on the basis of explanatory coherence. To overcome this limitation, Thagard has developed a computer program called ECHO⁷ that successfully implements the principles of explanatory coherence using connectionist techniques (Thagard, 1989, 1992).

⁷ ECHO stands for Explanatory Coherence by Harmany Optimization (Thagard, 1992). The term 'Harmany' is a tribute to Gilbert Harman and his early ideas regarding inference to the best explanation.

Briefly, in ECHO propositions corresponding to hypotheses and evidence are represented by nodes or neuron-like units, with links between them representing relations of coherence and incoherence. In accordance with Principle 1 (which states that propositions cohere and incohere equally), these links are symmetrical. Coherence relations (determined by Principles 2 and 3), are represented by excitatory connections and incoherence relations (specified by Principles 5 and 6), by inhibitory connections. Hence, if two propositions cohere because of explanatory relations holding between them, then the units representing these propositions are connected by an excitatory link. Conversely, if two propositions incohere because of relations of competition or contradiction, then the units representing these propositions are connected by an inhibitory link. ECHO implements Principle 4, concerning data priority, by establishing links from the units representing the data propositions to a special evidence unit with a constant activation of 1. Figure 5.5 depicts a simple connectionist network that demonstrates these properties.

When this network is run, activation flows from the special evidence unit to the evidence units (E1 and E2), and then to the units representing the explanatory hypotheses (H1 and H2). Because of the inhibitory link between H1 and H2, these units have to compete for the activation spreading from the evidence units and the activation of one will tend to suppress the other. In order to compute the acceptability of the competing explanatory hypotheses, ECHO uses a standard connectionist algorithm (see the Appendix at the end of this chapter for details), which repeatedly adjusts the activation of all the units in parallel over a specified number of cycles. In each cycle of activation adjustment, the activation of each unit is updated based on the activation of the units to which it is connected by excitatory and inhibitory links. This process is repeated until all the units have reached static or unchanging activation levels, indicating that the network has settled into a stable state. At this point, some units will remain activated, with a final activation above a threshold of 0, while other units will be deactivated (final activation < 0). These final activation levels serve to determine the acceptability of the propositions represented. Specifically, propositions represented by units that have positive activations are accepted, and those represented by units that have negative activations are rejected. The result of running the network depicted in Figure 5.5 is that H1, which demonstrates greater explanatory breadth, is

accepted (remains activated), and the competing explanatory hypothesis, H2, is rejected (deactivated).

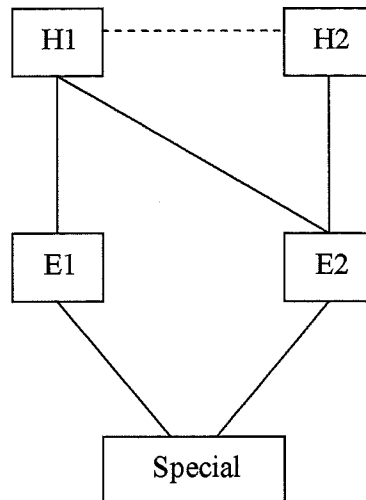


Figure 5.5 A simple ECHO network. Solid lines represent excitatory links and dashed lines represent inhibitory links. Evidence units are linked to a special unit with a constant activation of 1.

By demonstrating how the acceptability of competing hypotheses can be effectively computed on the basis of explanatory coherence considerations (specified by TEC's seven principles), ECHO can be seen to provide a concrete connectionist solution to the problem of theory evaluation. Thagard (1992) points out that repeated runs of the program have confirmed ECHO can successfully integrate the criteria of explanatory breadth, simplicity, and analogy emphasized by TEC, with activation accruing to units representing explanatory hypotheses that 1) explain more than their competitors, 2) are simpler, and 3) are analogous to other explanatory hypotheses. More substantively, Thagard has successfully applied ECHO to numerous cases from the history of science, including Lavoisier's argument for the oxygen theory and Darwin's argument for evolution by natural selection (Thagard, 1989, 1992, 1999; Eliasmith & Thagard, 1997; Nowak & Thagard, 1992). Collectively, these simulations lend strong support to the view that ECHO offers researchers a plausible model of theory evaluation that captures the strategies actually employed by scientists when faced with competing theoretical alternatives. With this comprehensive account of theory evaluation in

hand, therefore, we can return to the developmental findings of children's judgments on theory choice tasks and the question of parallels with scientists' judgments. Specifically, given Thagard's model, is it plausible to argue that children evaluate ideas on the basis of explanatory coherence considerations?

5.2.3 Do children apply principles of explanatory coherence?

In order to determine whether children evaluate their knowledge in line with the principles of explanatory coherence, we can identify the following four requirements that would need to be met (see also Thagard, 1992, Chapter 10):

1. Children possess coherent explanatory theories
2. They can appropriately differentiate and co-ordinate hypotheses and evidence
3. They are sensitive to the specific criteria that enter into judgments of explanatory coherence
4. They can use these criteria as a basis for theory evaluation.

1. Coherent theories From the evidence reviewed in this chapter, it can be argued that the first requirement is satisfied. In Vosniadou and Brewer's (1992, 1994) investigations of children's models of the earth and the day/night cycle, even young children were found to possess coherent explanatory models of the phenomena in question, which served to frame and inform their responses in detailed interviews. Similarly, Samarapungavan and Wiers' (1997) study of children's ideas about the origin of the species produced convincing evidence that children construct coherent explanatory frameworks that constrain the solutions they generate to a variety of biological problems. More generally, these investigations add to the substantial body of literature that ascribes theory-like qualities to intuitive knowledge (e.g., Brewer & Samarapungavan, 1991; Carey, 1985; Karmiloff-Smith, 1988; Karmiloff-Smith & Inhelder, 1974; McKloskey & Kargon, 1988; Slaughter & Gopnik, 1996; Smith, Carey & Wiser, 1985; Wellman, 1990; Wiser, 1988). Collectively then, this research supports the view that children possess systematic theoretical structures of the sort required for judgments of explanatory coherence.

2. Ability to differentiate hypotheses and evidence A second requirement that must be met before we can conclude that children apply principles of explanatory coherence, concerns their ability to differentiate hypotheses and evidence. While Thagard's model of theory evaluation specifically requires this ability, the question of whether children can appropriately distinguish theories from evidential support for or against those theories has been at the forefront of arguments regarding the development of scientific reasoning skills in children (and lay adults) more generally. For example, in the Introduction to this thesis, I highlighted claims by Deanna Kuhn (Kuhn, 1989; Kuhn, Amsel & O'Loughlin, 1988), that children and lay adults do not reason like scientists because they are unable to differentiate and co-ordinate hypotheses and evidence effectively. In contrast, Samarapungavan's (1992) investigation of children's performance on theory choice tasks reviewed in this chapter has led her to reject Kuhn's claim that children are deficient reasoners when it comes to coordinating theory with evidence. Specifically, her findings demonstrate that even first graders are competent in using both disconfirmatory and confirmatory evidence to choose between alternative theories, and can often justify their choices by explicit reference to key aspects of the theory-evidence relationship. In particular, Samarapungavan found that when justifying their choice of the empirically consistent theory, "... as many as 82% of the children investigated mentioned the evidence that undermined the rival theory as well as the evidence that supported the one they had selected" (Samarapungavan, 1992, pp.20-21).⁸

Consistent with Samarapungavan's findings, Sodian, Zaitchik and Carey (1991) have proposed that even young elementary school children can appropriately differentiate hypothetical beliefs from evidence. Specifically, these researchers presented first and second graders with two competing hypotheses and asked them to select an empirical test that would allow them to decide between the hypotheses. They found that the majority of the children correctly selected a conclusive test and were able to

⁸ Further, she proposes that Kuhn's (1989) findings may be more an outcome of the specific tasks employed to assess reasoning ability, than the inability of children and lay adults per se to differentiate and co-ordinate theories and evidence. In particular, Samarapungavan points out that Kuhn's subjects were required to evaluate the cumulative co-variation between possible causes and effects across several test tasks, a complex form of theory-evidence relationship that goes beyond an ability to differentiate and co-ordinate theories and evidence (Samarapungavan, 1992).

distinguish it from an inconclusive one. Moreover, in a further task that presented children with a genuine scientific problem, children were found to spontaneously generate strategies for gathering evidence to decide between alternative hypotheses (Sodian, Zaitchik & Carey, 1991).

3. Sensitivity to criteria relevant to explanatory coherence judgments What about the third requirement listed above, namely that children should demonstrate an appreciation of the factors relevant to assessments of explanatory coherence? In addition to demonstrating that children possess robust explanatory models of natural world phenomena, Vosniadou and Brewer also uncovered an appreciation of some of the factors identified by Thagard (1992) as relevant to assessments of theory quality. In particular, Vosniadou and Brewer (1994) found children's models of astronomical phenomena were internally consistent and demonstrated a commitment to principles of simplicity. Relatedly, Samarapungavan and Wiers (1997), draw on Thagard's proposals to argue that children's ideas about speciation comprise "... an internally consistent interrelated set of core beliefs" (Samarapungavan & Wiers, 1997, p.167). They conclude that their findings support those of other researchers such as Vosniadou and Brewer, indicating that children's conceptual systems demonstrate properties of explanatory coherence (Samarapungavan & Wiers, 1997).

4. Use explanatory coherence criteria as a basis for theory evaluation On the basis of the evidence reviewed, it would appear that the first three requirements for determining whether children can reason on the basis of explanatory coherence considerations are satisfied. Detailed reconstructions of children's explanatory frameworks across a number of domains indicate children have recourse to coherent theories that demonstrate essential properties of explanatory coherence and which can be appropriately employed by them in their everyday problem solving. In addition to these findings, it is suggested that Samarapungavan's (1992) study provides direct evidence for requirement 4, by showing that children can also use criteria relevant to explanatory coherence judgments as a basis for theory selection.

In particular, Samarapungavan's finding that children selected theories that explained a greater range of observations can be seen to conform directly to Thagard's criteria of explanatory breadth embodied in TEC, and can be analysed as a coherence problem

using the ECHO network displayed in Figure 5.6. On ECHO computations of explanatory coherence, the more a theory explains, the greater its coherence and therefore the more likely it will be preferred over its competitor. In this network, which depicts the ‘range of explanation’ condition on the chemistry task given to children, T2 can be seen to demonstrate greater explanatory coherence than the competing theory T1, by explaining more of the evidence.

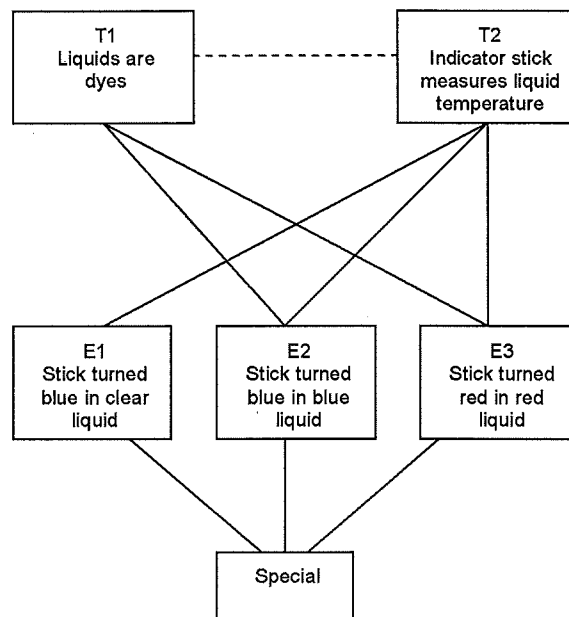


Figure 5.6 An ECHO network depicting the ‘range of explanation’ condition on the chemistry task.

Similarly, Samarapungavan’s finding that older children showed a systematic preference for theories that did not rely on ad hoc hypotheses can also be modelled using ECHO (Figure 5.7). Specifically, the notion of simplicity embodied in Thagard’s theory of explanatory coherence is one of non-ad hocness, according to which the degree of coherence of a hypothesis with the evidence and with its co-hypotheses is inversely proportional to the number of co-hypotheses required. Implementing this criterion in ECHO, results in a preference for theories that make fewer ad hoc assumptions. In this network, which depicts the ‘non-ad hocness’ condition on the chemistry task, T1 introduces an auxiliary hypothesis, A1, to deal

with the outstanding observation that the indicator stick immersed in the colourless liquid turns blue. T2 is more likely to be preferred over T1 on the basis of its greater explanatory coherence, because it provides a simpler explanation of the evidence that does not rely on any auxiliary hypotheses. Therefore, in both cases children's judgments relating to these criteria are consistent with predictions made by Thagard's model of theory evaluation.⁹

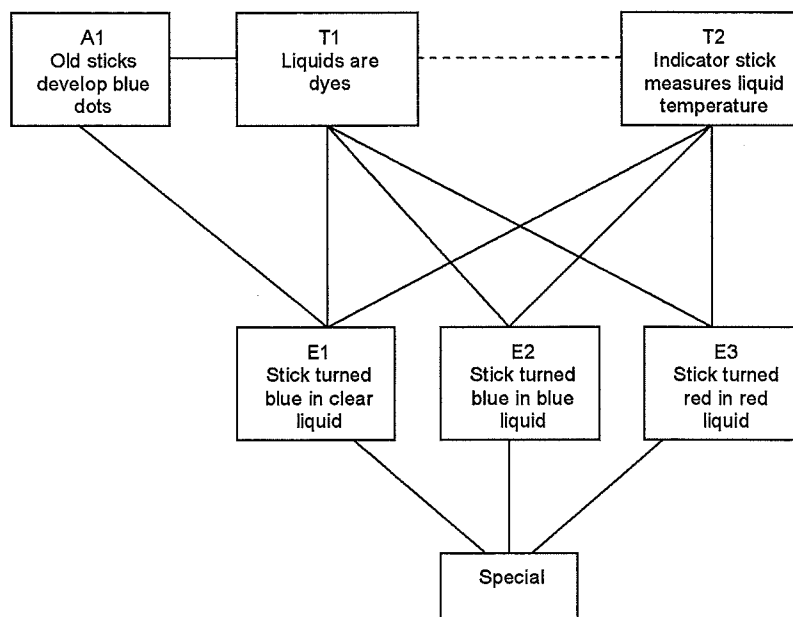


Figure 5.7 An ECHO network depicting the 'non-ad hocness' condition on the chemistry task.

⁹ What about the criterion of analogy, which Thagard argues can contribute to explanatory power in comparative assessments of competing theories? Samarapungavan (1992) does not directly investigate this criterion and therefore provides no insights concerning its possible use by children in theory choice contexts. However, indirect evidence that children may have some appreciation of this factor is provided by studies of analogical reasoning in children (e.g., Goswami & Brown, 1990a, 1990b) that show the ability to reason by analogy is present very early on and provides a building block for subsequent learning. Moreover, Thagard (2000) has argued that the commonsense assumption for the existence of other minds at the heart of our folk psychological theorizing can be construed as an explanatory coherence problem with analogy (to our own minds) playing a central role. Finally, it is important to note that Thagard's (1992) analysis of the role of explanatory coherence in scientific revolutions, found that analogy was only minimally important to scientists' arguments (i.e., of the seven revolutions examined, analogy played a small role in only one, namely Darwin's argument for evolution by natural selection).

In addition, findings of children's preference for theories that were empirically and logically consistent (the two additional criteria investigated by Samarapungavan, 1992), would also be predicted by Thagard's model. Firstly, concerning the issue of empirical consistency, this is addressed in TEC by Principle 4, data priority, which allocates the results of observation or experiment a degree of independent acceptability. As Thagard makes clear, the inclusion of this principle means that TEC is not a "pure coherence theory" (Thagard, 2000, p.43), but instead acknowledges that empirical consistency features in judgments of the explanatory worth of competing theories. Moreover, the way in which this principle is implemented in ECHO, with activation spreading in a non-symmetrical fashion from the evidence units to the hypotheses units via excitatory and inhibitory links, reveals how 'consistency with evidence' acts as a crucial determinant of acceptability.¹⁰

Secondly, regarding children's preference for theories that were free of internal contradictions or inconsistencies, the local relations of coherence and incoherence between propositions that are established by TEC's principles of Explanation and Contradiction/Competition, determine that ECHO will demonstrate a natural preference for theories that are internally consistent over those that contain contradictory propositions. More specifically, contradictions contribute to judgments of explanatory coherence when the competing theory is either internally contradictory or explains "negative evidence" (Thagard, 1992). In discussing this point, Thagard uses the example of ECHO's preference for the oxygen theory over the phlogiston theory, which demonstrated internal contradictions by assuming that phlogiston could have both positive and negative weight.

Finally, Samarapungavan's findings that the plausibility of the theories presented (determined by their consistency/inconsistency with children's prior knowledge) had an impact on children's evaluative strategies, can also be interpreted within Thagard's model of theory evaluation, where the presence of higher-level explanations can affect explanatory coherence judgments. In particular, principle 2(a) of TEC determines that a hypothesis is more coherent if it is itself explained. Correspondingly, in ECHO a

¹⁰ Related to this point, Haig has recently suggested that the criterion of explanatory breadth in TEC can be understood as a non-predictive measure of empirical adequacy (Haig, 2002).

hypothesis that is explained by a higher-level hypothesis gains activation from this explanatory relation. Conversely, a hypothesis that is inconsistent with a higher-level hypothesis would tend to be disadvantaged by the inhibitory link that is established on the basis of incoherence relations. Likewise, Samarapungavan (1992) found that children were more likely to choose the 'correct' theory in terms of the criteria being examined when it did not contradict their existing explanatory frameworks.

In summary, on the basis of Samarapungavan's direct study of children's theory selection strategies, it is plausible to argue that requirement four is also met. Samarapungavan's findings show that children can appropriately evaluate competing theories, using criteria that Thagard's explanatory coherence model identifies as pivotal for scientific theory evaluation.

5.3 The pervasiveness of coherence-based inference

In this chapter I have developed the proposal that the abductive framework offers a source of fruitful ideas about children's knowledge construction strategies by investigating the possibility of significant relations between children and scientists in the context of theory evaluation. With this aim in mind, I began by reviewing a body of research detailing children's theories of the natural world that points to substantial correspondences between the criteria used by scientists and children to evaluate knowledge. Following this review, I proposed that the abductive framework, with its endorsement of a precise computational model of theory evaluation, could illuminate these correspondences by showing how the criteria identified by developmental researchers enter into judgments of explanatory hypotheses on the basis of explanatory coherence.

In particular, this model of theory evaluation indicates that children can employ explanatory coherence considerations as a basis for choosing between competing theories. Children demonstrate a preference for theories that explain more facts, require fewer ad hoc hypotheses to achieve this end, and are consistent both internally and with the evidence, indicating the likelihood that children can adopt a multi-

criteria perspective on theory evaluation in a similar manner to scientists. While further research is needed to determine whether children can spontaneously integrate these criteria into more complex judgments of explanatory coherence, a recent proposal by Thagard (2000) that coherence-based reasoning is pervasive in human thought and action, lends support to the view that children readily apply complex coherence considerations in their everyday reasoning.¹¹ Specifically, Thagard (2000) builds on existing connectionist applications of coherence to the problems of explanatory inference (Thagard, 1989, 1992), analogical reasoning (Holyoak & Thagard, 1989a, 1995), and decision-making (Thagard & Millgram, 1995; Millgram & Thagard, 1996), to argue that ideas about coherence are central to solving the general puzzle of how we make sense of the world. According to Thagard, “making sense” is best conceived as a coherence problem that involves “. . . fitting something puzzling into a coherent pattern of mental representations that include concepts, beliefs, goals, and actions” (Thagard, 2000, xi). In this view, many of the inferences we make in scientific and everyday contexts, do not take the form of stepwise linear reasoning epitomized by the canons of formal logic. Rather, Thagard argues they involve holistic judgments, in which large sets of elements are simultaneously assessed in order to determine how they fit together into a satisfying whole (Thagard, 2000). Given the proposed centrality of coherence mechanisms in both scientific and everyday thought, then, it would seem plausible that further investigations may well reveal children’s capacity for integrating coherence criteria into complex judgments of explanatory coherence, in a way that parallels ECHO implementations of multi-criteria theory evaluation in science.

¹¹ It is also worth noting here that in an open peer commentary on Thagard’s theory of explanatory coherence (Thagard, 1989), Carl Bereiter (Bereiter & Scardamalia, 1989) has reported that elementary school children can use explanatory coherence reasoning effectively in instructional contexts, and remarks on the speed at which the students caught on to the ‘logic’ of ECHO, viewing it as natural and reasonable.

5.4 Appendix

Equation used to update activation in ECHO network (refer Thagard, 1992, p.101; see also Thagard and Verbeurgt, 1998):

On each cycle the activation of a unit j , a_j , is updated according to the following equation:

$$a_j(t + 1) = a_j(t)(1 - d) + net_j(max - a_j(t)) \text{ if } net_j > 0, net_j(a_j(t) - min) \text{ otherwise}$$

Here d is a decay parameter (say .05) that decrements each unit at every cycle, min is minimum activation (-1), max is maximum action (1). Based on the weight w_{ij} between each unit i and j , we can calculate net_j , the net input to a unit, by:

$$net_j = \sum_i w_{ij} a_i(t).$$

Weights can be positive (typically .04) representing excitatory links, or negative (typically -.06), representing inhibitory links.

Chapter 6

Speculating on the Cognitive Origins of Science

In Chapters 4 and 5, I argued that a methodological perspective on the child-as-scientist debate could inform investigations of the process of knowledge acquisition in childhood by highlighting some significant parallels between the abductive inferential strategies employed by children and scientists. Having identified these parallels, this chapter turns to speculate on the cognitive origins of science and, in particular, the claim by proponents of the theory theory to have solved the evolutionary puzzle of our cognitive capacity for science.

Accordingly, in Section 6.1 I begin by detailing the evolutionary story underpinning the theory theory. Section 6.2 then surveys a range of criticisms levelled at Gopnik and Meltzoff's (1997) claim to have identified science's cognitive foundations in early childhood learning. Having discounted the theory theory's evolutionary speculations, attention turns to an alternative proposal put forward by Mithen (2002) that the key components of scientific reasoning emerged in a gradual fashion over the course of human evolution. Finally, Section 6.3 briefly considers some recent indications of continuities between scientific and pre-scientific reasoning practices (Carruthers, 2002b) that hold particular relevance for the methodological parallels developed in this thesis, and concludes that these continuities offer a promising avenue for future investigations of the cognitive origins of science.

6.1 The theory theory's solution to an evolutionary puzzle

While the last two chapters have developed a methodological perspective on the child-as-scientist debate as an alternative to the theory theory, the focus of this chapter shifts to speculations about the cognitive origins of science. What makes scientific inquiry possible? Are the foundations of science to be found in childhood learning? To what extent are there continuities across the reasoning strategies employed in pre-scientific cultures and those taken to be central to scientific inquiry? Although attempts to answer these questions will be necessarily speculative, the issue of science's cognitive basis can be seen to have relevance for investigations of child-scientist parallels in at least two fundamental respects.

Firstly, part of my rationale for adopting a science-as-method view and reformulating child-scientist parallels along methodological lines (refer Chapter 2), appealed to the evolutionary naturalist insights of Hooker (1987, 1989). At a fundamental level, this endorsement of an evolutionary approach to epistemological issues entails a commitment to viewing ourselves, including our minds, as the products of evolution. Adopting an epistemology that is conceptually integrated with scientific knowledge of our evolved status, in turn leads one to expect fundamental continuities in knowledge seeking strategies across species, and in the case of humans, across scientific and everyday cognition. Moreover, this expectation forms the basis of any claim by proponents of the child-as-scientist analogy for substantive parallels between childhood cognition and theory change in science. Therefore, incorporating some preliminary speculations about the cognitive foundations of science into my analysis of child-scientist parallels is both consistent with the broad epistemological perspective informing this thesis, and relevant to assumptions of continuities in inferential practices that lie at the heart of the child-as-scientist debate.

Secondly, in addition to recognizing the significance of our evolutionary history when attempting to develop child-scientist parallels, a further reason for considering the origins of our scientific abilities relates to claims embodied in the theory theory itself. In Chapter 1, I indicated that an evolutionary explanation for the existence of substantive parallels between cognitive development and scientific change forms a

core component of Gopnik and Meltzoff's (1997) theory theory. In particular, their argument for the presence of common theory formation capacities rests on a specific evolutionary story regarding the adaptive function of theory building devices in early childhood learning contexts. Moreover, Gopnik and Meltzoff claim that this account of child-scientist parallels solves an "interesting evolutionary puzzle" regarding our cognitive capacity for science:

Where did the particularly powerful and flexible cognitive devices of science come from? After all, we have only been doing science in an organized way for the last 500 years or so; presumably they didn't evolve so that we could do that. We suggest that many of these cognitive devices are involved in the staggering amount of learning that goes on in infancy and childhood. Indeed, we might tell the evolutionary story that these devices evolved to allow human children, in particular, to learn. (Gopnik & Meltzoff, 1997, p.18)

According to this view, the "cognitive devices" underpinning scientific inquiry are seen to constitute a basic design feature of human minds that evolved to facilitate essential learning about the causal structure of the world. The puzzle of the origin of our ability to engage in scientific endeavour is therefore solved by adopting a developmental approach, and focusing on the natural links between scientific change and knowledge acquisition in childhood.

Have Gopnik and Meltzoff (1997) provided a solution to the puzzle of our capacity to undertake science? With this question in mind, the following section reviews the evolutionary speculations that form part of Gopnik and Meltzoff's proposals in more depth by attempting to trace the emergence of a specific evolutionary component to the theory theory across a series of successive formulations.

6.1.1 Science is a 'by-product' of childhood

An initial indication that there might be an evolutionary reason for the existence of child-scientist parallels appears in an early defence of the theory theory by Alison Gopnik and Henry Wellman (Gopnik & Wellman, 1992). In this paper, the authors conclude their argument with the provocative suggestion that "[s]cience . . . might be a sort of spandrel, parasitic on cognitive development itself" (Gopnik & Wellman, 1992, p.168). While not elaborating on this speculation in any detail, they introduce the possibility that the processes of scientific inquiry are epiphenomenal consequences of more fundamental learning processes seen in childhood. By doing so, the authors also indicate that further examination of the capacities for reasoning about the causal structure of the world underlying scientific inquiry may reveal a surprisingly close relationship with the inferential strategies involved in cognitive development.

The possibility that science as a relatively recent cultural invention is a '*by-product*' of the commonsense learning that occurs in infancy and early childhood is further developed in a subsequent chapter by the authors (Gopnik & Wellman, 1994). In particular, this evolutionary speculation can be seen to be tied to their increasing interest in a cognitive characterization of science, both as a means of clarifying the theory theory, and as a way of distinguishing it from competing accounts:

. . . scientific theory change and conceptual change in childhood are both the product of human minds trying to understand the world around them. Scientific change must centrally involve some human cognitive capacity.
(Gopnik & Wellman, 1994, p.258)

Having established the possibility of a shared cognitive basis between scientific and childhood thought, Gopnik and Wellman (1994) subsequently begin to specify the precise nature of this proposed connection. Contrary to the assumed direction of insights from science to childhood learning, the authors argue that the opposite may in fact be true:

. . . the similarities are better captured by thinking of scientists as big children, rather than thinking of children as little scientists. The progress of science, we believe, reflects certain fundamental processes of conceptual change that are first seen in very young children. (Gopnik & Wellman, 1994, p.259)

The motivation for this ‘reflexive’ move by the authors is clearly not tied to pragmatic concerns regarding the accessibility of children as research subjects, or the (comparative) ease of studying childhood cognitive development in contrast to historical cases of scientific change.¹ Rather, Gopnik and Meltzoff’s suggestion is that the cognitive capacities seen in childhood are *epistemologically prior* to those operating in science. Moreover, they speculate that any explanation of this relationship is likely to involve an ultimate account of these capacities, in terms of the function they served in facilitating childhood learning in our hunter-gatherer past (Gopnik & Wellman, 1994).

In addition to highlighting a possible evolutionary foundation for the theory theory, Gopnik and Wellman (1994) also speculate about the form of this evolutionary endowment by contrasting the theory theory with a modularity account of children’s developing understanding of the mind. On the modularity view (e.g., Leslie, 1987), the young child’s emerging folk psychology is seen to comprise an evolved module or series of modules dedicated to belief-desire reasoning, which come ‘on-line’ at specific points in development. In contrast to this speculation, Gopnik and Wellman suggest that the picture of development emerging from empirical investigations speaks against such a modular view:

It is much more difficult to see how evolution would have selected for a series of representational systems, each maturing separately only to be replaced by another. (Gopnik & Wellman, 1994, p.284)

¹ Such ‘pragmatic’ motivations have been identified in both Kuhn’s and Piaget’s investigations of the relations between childhood and science (see Levine, 2000, for a discussion).

Instead, the authors argue for the evolution of a domain-general theorizing capacity, which, in combination with some domain-specific initial knowledge or “starting-state theories”, is seen to be capable of producing the dynamic picture witnessed in development:

What seems more plausible is that evolution selected for a cognitive capacity to revise concepts on the basis of evidence, that is, a theory-making ability. (Gopnik & Wellman, 1994, p.284)

6.1.2 Evolved mechanisms and scientific progress

The idea that evolution has endowed us with a general and flexible theorizing capability is further elaborated in Gopnik’s (1996b) formulation of the theory theory. In this target paper, and in her reply to commentators, Gopnik emphasizes the bi-directional nature of the analogy, and sets about developing the claim that the epistemological success of science can be explained via its cognitive links to childhood learning. In particular, she argues that the most critical feature of science requiring explanation is its ability to “get things right”, and suggests the success of science may be largely due to the exploitation of powerful learning mechanisms which evolved to facilitate knowledge acquisition in early childhood.

In order to support this claim, Gopnik develops the following argument. Firstly, in keeping with Gopnik and Wellman (1994), she advocates a cognitive perspective on science, according to which the focus of inquiry is on the cognitive abilities employed by scientists in their knowledge seeking endeavours:

... scientists must be using some cognitive abilities to produce new scientific theories and to recognize their truth when they are produced by others. Scientists have the same brains as other human beings, and they use those brains, however assisted by culture, to develop knowledge about the world. Ultimately, the sociology of science must consist of a set of individual decisions by individual humans to produce or accept theories. (Gopnik, 1996b, p.487)

Secondly, in conjunction with this perspective, Gopnik endorses what she takes to be the basic idea underpinning cognitive science that evolution has endowed us with a variety of devices for obtaining an accurate representation of the world. Thirdly, she suggests that scientific inquiry capitalizes on these devices in its pursuit of the truth:

A cognitive scientist would say that evolution constructed truth-finding cognitive processes. Science employs a particularly powerful and flexible set of these cognitive abilities. Science uses a set of representations and rules that are particularly well-suited to uncovering the truth about the world. Science gets it right because it uses psychological devices that were designed by evolution precisely to get things right. (Gopnik, 1996b, p.489)

By emphasizing the continuity of scientific and everyday cognition, and the dependence of human cognitive abilities on our evolutionary circumstances, Gopnik acknowledges her debt to the naturalistic epistemological tradition of Quine and others (e.g., Quine & Ullian, 1970²). However, Gopnik also makes it clear that she is advocating a developmental version of the naturalistic thesis that indicates the possibility of a unique link between science and cognitive development. In particular, she argues that cognitive science claims for “truth finding” devices that are exploited by science, raise obvious questions about the origin of these devices. Her solution to this evolutionary puzzle is that they are present in the essential early learning of infancy and childhood, and that an account of their origins would reflect this developmental context.

In making these claims, Gopnik highlights three characteristic features of humans when compared to other species: our behavioural plasticity, adaptive flexibility, and the relative lack of specialized cognitive abilities at birth, as reflected in a long period of immaturity (Gopnik, 1996b). Gopnik suggests that given these distinctive features of humans, the provisioning of a general theory building ability in the young of our species would make sound evolutionary sense:

² For Quine’s classic presentation of the naturalistic epistemology thesis, see his essay *Epistemology naturalized* in Quine (1969, pp.69-90).

Equipping human children with particularly powerful and flexible cognitive devices, devices that are good at constructing accurate representations of new and unexpected worlds, might be an important part of this evolutionary strategy. We might indeed think of childhood as a period when many of the requirements for survival are suspended, so that children can concentrate on acquiring a veridical picture of the particular physical and social world in which they find themselves. Once they know where they are, as it were, they can figure out what to do. (Gopnik, 1996b, p.490)

On this view, then, attention to the evolutionary origins of theory building is taken to support Gopnik's claim that the analogy "cuts both ways", and that childhood cognition has explanatory import for understanding science. In particular, it is argued that science co-opts the natural learning mechanisms that evolved to facilitate learning in young children, and puts them to work in a "culturally-organized way". This suggests that any adequate account of scientific progress will need to take account of these evolved mechanisms and their origins in early commonsense learning. As Gopnik remarks:

To explain scientific theory change we may need to talk about culture and society, but we will miss something important if we fail to see the link to natural learning mechanisms. (Gopnik, 1996b, p.493)

Finally, in Gopnik and Meltzoff's (1997) defence of the theory theory, these evolutionary speculations are packaged together and explicitly presented as a core component of the theory theory account of child-scientist parallels. The authors claim firstly that the basic cognitive apparatus of science – the ability to generate and revise theories – is part of our evolutionary endowment. Secondly, they suggest that these theorizing capacities evolved specifically to facilitate knowledge acquisition in infancy and early childhood. Thirdly, they argue that science targets and exploits these innate capacities and, as a result, is largely successful in its attempts to construct an accurate representation of reality.

Taken together, these three claims can be seen to comprise a particular evolutionary story about the origins of theory building which serves as a foundation for Gopnik and Meltzoff's more general proposal that "... cognitive development in childhood may be much like scientific theory change" (Gopnik & Meltzoff, 1997, p.29). As they point out, the task facing theory theory researchers is not to demonstrate that young children "do science", but to argue for significant deep-structural parallels between the cognitive processes supporting scientific endeavour and those engaged in the bulk of cognitive development. Moreover, in making their case for the theory theory, these speculations about the origins of human theory building are seen to solve the puzzle of our human capacity for science.

6.2 What makes scientific inquiry possible?

6.2.1 Have Gopnik and Meltzoff solved the evolutionary puzzle?

In the preceding section, I have attempted to trace the emergence of a distinct evolutionary component to the theory theory, beginning with the brief statement found in Gopnik and Wellman (1992), and culminating in the evolutionary story presented by Gopnik and Meltzoff (1997). However, while these authors clearly rely on evolutionary considerations as one form of evidence for the theory theory, a number of commentators have questioned their claim to have identified the cognitive basis of science. These critiques can be seen to fall into three broad categories: assessments of Gopnik and Meltzoff's evolutionary account as a 'just-so' story (Fine, 1996; Stich & Nichols, 1998); criticisms of the authors' commitment to "truth-tropic cognition" and the idea that evolution has selected for truth-finding cognitive processes (Downes, 1999); and claims that the theory theory is unable to account for the emergence of science in human history, or "the 1492 problem" (Giere, 1996; Faucher et al., 2002). I briefly consider each in turn.

Theory building as an evolutionary adaptation: A ‘just-so’ story? An initial criticism of the proposal that theory building constitutes an evolutionary adaptation, concerns the lack of appropriate evidence. For example, Fine (1996) takes Gopnik’s (1996b) account of the innate basis of scientific reasoning to be a “thin tale”, which fails to meet the evidentiary standards of evolutionary biology. More specifically, he points out that Gopnik’s claims for innate theory-formation devices are unconstrained by any details of the environments in which these devices initially emerged and the selective pressures known or thought to have been operating within them. In the absence of any evidence for “. . . differential variation and selective fitness with respect to specific local environments” (Fine, 1996, p.535), Fine argues that Gopnik is unable to solve the very problem she claims to have solved, namely how science “gets things right”. As indicated in Section 6.1.2, Gopnik identifies scientific progress as the feature of science that is most in need of explanation and argues the reason scientists manage to get at the truth is that they exploit theory revision mechanisms designed by natural selection to facilitate early learning. Yet, as Fine remarks, without providing details of the selection and development of a theory building trait in human children, Gopnik is at a loss to explain the presence of these mechanisms and, therefore, the success of science.

Similarly, Stich and Nichols (1998) suggest that while Gopnik and Meltzoff’s (1997) evolutionary speculations are broadly compatible with a range of facts to be explained, other ‘just-so’ stories could be constructed that would be equally consistent with the data. More specifically, they point out that the theory theory’s ‘solution’ in its current form actually raises a further evolutionary puzzle about the function of the theory building mechanisms underpinning scientific inquiry. That is, Gopnik and Meltzoff’s commitment to “theories all the way down” is at odds with their claim that the adaptive virtue of a theory building capacity is to enable children to make sense of variable aspects of the environment, given that much of the early learning they discuss involves stable, highly *invariant* features of the world. In Stich and Nichols’ view, this “tension” in Gopnik and Meltzoff’s account reveals a lack of clarity about the

sorts of information processing problems the theory formation system was designed to solve and leaves the question of the origins of scientific cognition open to debate.³

Cognition, fitness, and truth Further criticisms of Gopnik and Meltzoff's evolutionary story are found in Downes (1999), who argues that the theory rests on questionable assumptions about the nature and origin of human cognition. In particular, he suggests that Gopnik and Meltzoff's claims of child-scientist parallels are inextricably linked to their particular view of cognition as strongly veridical or "truth-tropic". As Downes remarks, Gopnik and Meltzoff take natural selection to be responsible for the emergence of truth-gaining cognitive processes in humans, and propose that science utilizes a particularly powerful set of these evolved cognitive capacities in its "pursuit of the truth". This picture of truth-tropic cognition and science is then linked to children's cognitive development by arguing that the reason children's and scientists' cognitive processes look similar is because of their evolutionary history – they evolved to allow human children to achieve truthful representations of the world.

According to Downes (1999), however, these presuppositions about the veridicality of human cognition and its evolutionary origins are problematic for understanding scientific development in at least two respects. Firstly, he questions whether cognition and truth-tropic cognition are in fact the same thing, and cautions against conflating naturalistic hypotheses about actual mental processing with normative ones about optimal or ideally rational reasoning. In the case of science, he argues that an uncritical commitment to the idea that good science is 'truth-attaining' runs the risk of conflating the success of science "construed as the collection of practices, combined

³ One possible story, suggested by Stich and Nichols (1998), is that scientific reasoning may be a 'by-product' of later childhood learning rather than infant cognition as Gopnik and Meltzoff (1997) maintain. More generally, it is not clear on Gopnik and Meltzoff's account why a capacity for theory building is functionally tied to the sorts of information-processing problems experienced in childhood at all. Despite their claims that comparative studies indicate the plausibility of powerful early learning mechanisms in humans, it is conceivable that a capacity for forming and revising theories emerged because of its role in solving any of a number of adaptive problems requiring imaginative thinking that are encountered in *adulthood* e.g., problems of parenting, hunting animals, gathering plant foods, selecting a good mate, etc. Therefore, even if young children demonstrate a facility for forming and revising theories, this does not necessarily indicate the evolutionary function of theory building is to enable children "to get things right". On this point, see Carruthers' (2002a) recent proposal that the forms of pretend play witnessed in early childhood have not evolved to service children's needs, but have been selected for because of their role in enhancing *adult* creativity.

throughout history” with individual scientists’ cognitive practices as truth-tropic. Secondly, Downes suggests that Gopnik and Meltzoff’s claim for the evolution of veridical cognitive mechanisms rests on a questionable relation between ‘fitness’ and ‘truth’. Drawing on the work of philosophers advocating a pragmatic approach to epistemology (e.g., Stich, 1990), Downes cites examples in which fitness and truth are not aligned to support his proposal that “truth is separable from cognitive success”, and that despite Gopnik and Meltzoff’s claims, “truth-tropic cognition is not likely to be selected for” (Downes, 1999, p.575).

“The 1492 problem” A final argument against the theory theory’s proposed solution to the evolutionary puzzle concerns its inability to account for science’s historical development. For example, Giere (1996) questions why modern science took so long to make its appearance in human history if the cognitive capacities sufficient for undertaking scientific inquiry emerged in the Pleistocene. More particularly, he argues that scientific activity as we understand it was absent in 1492, but given the short time span between now and then we cannot attribute its subsequent appearance to evolved cognitive capacities. Therefore, Giere concludes that something else must have been responsible for the development of science, something that is missing from Gopnik’s (1996b) ‘solution’. Giere then proceeds to identify some of the elements he thinks are implicated in the scientific revolution, including the development of instrumentation, symbolic notations, printed materials, and experimental methods, and all of which are indispensable “for understanding the workings of modern science” (Giere, 1996, p.538).

In reply, Gopnik (1996b) suggests that the apparent gulf between pre-scientific and scientific cognition, and hence the difficulty of explaining the emergence of science, is resolved if the focus shifts from adult reasoning to the commonsense theorizing activities of young children. Having redirected attention to early childhood reasoning, she argues that in 1492, due to increased leisure and a range of technological and social advances, a raft of new evidence relevant to solving fundamental problems about the causal structure of the world was generated. This availability of evidence in turn led the innate theory formation devices of childhood to be ‘reactivated’ in the evolutionary novel context of adult science (Gopnik, 1996b).

However, as Faucher et al. (2002) have recently pointed out, this attempt to save the theory theory's solution fails because it is unable to explain why science did not make its appearance much earlier in China. In particular, they argue that well before 1492, the social and technological factors cited by Gopnik as crucial to the production of new evidence, were present in Chinese society. Given that "science as we know it did not emerge in China" (Faucher et al., 2002, p.340), Faucher et al. suggest that other critical factors must have been responsible for its subsequent appearance in the West. Having rejected the theory theory's solution to the evolutionary puzzle, they go on to identify a range of additional components needed to solve the 1492 problem that focus on the social and cultural transmission of norms, theories, and theory revision mechanisms. Moreover, Faucher et al. argue that these factors are important for attempts to understand scientific cognition.

In summary, the above criticisms can be seen to raise a number of questions about the utility of Gopnik and Meltzoff's evolutionary story for understanding the cognitive basis of science. The 'just-so' nature of their account, its reliance on presuppositions of truth-tropic cognitive processes and their evolution by natural selection, together with its inability to explain the emergence of science in human history, all undermine Gopnik and Meltzoff's claim to have identified the locus of scientific reasoning in early childhood cognition. If we reject the theory theory's solution, however, then we are left with the evolutionary puzzle of our ability to engage in science. In this respect, a recent collection of papers targeting the cognitive foundations of scientific inquiry (Carruthers, Stich & Siegal, 2002), can be seen to offer a useful platform for some initial speculations. Accordingly, in what follows, I first review evidence for the proposal that science demonstrates a piecemeal evolutionary history, before going on to sketch some ideas regarding continuities between scientific and pre-scientific cognition that centre on the capacity for abductive inference.

6.2.2 The emergence of scientific abilities in human evolution

In Chapter 2, I suggested that adopting a view of science-as-method indicates the likelihood of continuities in knowledge seeking strategies not only across science and our everyday cognition, but also across species more generally. In a similar vein, Robin Dunbar (Dunbar, 1996) has argued that when we shift from viewing science as a body of theory to recognizing it as a process or method of inquiry, scientific activity is revealed as “. . . a highly formalized version of something very basic to life, namely the business of learning about regularities in the world” (Dunbar, 1996, p.58). More particularly, he cites a range of evidence drawn from anthropology, psychology, and behavioural biology, to propose that two fundamental components of scientific activity – classification and causal inference – are commonplace in everyday human reasoning, and are also “. . . key feature[s] in the lives of most birds and mammals” (Dunbar, 1996, p.58). He concludes that science is not specific to modern Western culture or even peculiar to human beings, but rather is “. . . a ‘natural’ approach to the physical world . . . characteristic of all higher organisms” (Dunbar, 1996, p.77).

One implication of Dunbar’s ‘continuity’ claims for attempts to trace the cognitive origins of science is that certain fundamental aspects of scientific reasoning are likely to demonstrate a long evolutionary history. Recently, Steven Mithen (Mithen, 2002) has developed this suggestion further, arguing that our cognitive capacity for science has a biological basis that emerged gradually via a series of independent evolutions over a period of at least five million years. In order to support this proposal, Mithen begins by identifying what he takes to be the critical properties of science, and argues that the following components are essential to contemporary scientific practice:

- i) In-depth observation of the natural world
- ii) The ability to generate and test hypotheses
- iii) A focus on causation
- iv) The use of tools (such as external notational devices) to extend referential reach and facilitate problem solving
- v) The accumulation of knowledge and understanding over time
- vi) The use of metaphor and analogy.

Having established a working definition of science, Mithen turns to the fossil and archaeological records in an attempt to uncover evidence relating to the emergence of these components of scientific thought over the course of human evolutionary history. For expository purposes, he divides this 'history' into four roughly consecutive categories, beginning with the Early hominines (4.5 – c.1.8 mya), then moving to consider various species of Early humans (1.6 mya – 300,000 years ago), the Neanderthals (250,000 – 28,000 years ago), and finally *Homo sapiens* (beginning 130,000 years ago), which are further subdivided into hunter-gatherers of the late Pleistocene and early Holocene farmers.

Table 6.1 provides a summary of Mithen's (2002) account of the emergence of key elements of scientific inquiry in human evolution, drawn from his discussion of the anatomy and activities of these four groups of human ancestors and relatives. Concerning the first category of early hominines (australopithecines and early species of *Homo*), Mithen suggests that members of this group are likely to have demonstrated two fundamental components of scientific thought: a capacity for in-depth observation of the natural world, and a facility for generating and testing hypotheses. Evidence for the early emergence of these basic scientific abilities is argued to be found in the manufacture of stone tools during this period, which Mithen suggests would have required a form of hypothesis testing. In addition, he highlights the foraging activities of early hominines and suggests that after 2 mya these activities appear to have included systematic searching for animal carcasses, which would indicate the presence of rudimentary predictive capabilities. In making these claims, Mithen points out that the scientific skills he attributes to early hominines may not differ markedly from those found in the great apes, thereby aligning himself with Dunbar's (1996) general position that 'basic science' is characteristic of many animal species. However, contrary to Dunbar's claims that causal inference is a key component of these basic scientific abilities, Mithen stops short of attributing causal reasoning to early hominines, suggesting instead that such a facility makes a later appearance in early humans.

Table 6.1 The emergence of key elements of science in human evolution, compiled from Mithen (2002). The complete cognitive foundations for scientific activity are argued to be in place in hunter-gatherers of the late Pleistocene, with economic and social conditions conducive to the development of new bodies of knowledge subsequently appearing in early farming communities.

KEY ELEMENTS OF SCIENCE	Early hominines	Early humans	Neanderthals	Homo sapiens (i) (late Pleistocene hunter-gatherers)	Homo sapiens (ii) (early Holocene farmers)
Detailed observation of natural world	•	•	•	•	•
Hypothesis generation & testing	•	•	•	•	•
Concern with causation		• confined to theory of mind	• theory of mind only?	•	•
Use of tools to extend cognition			?	•	•
Accumulation of knowledge				•	•
Use of metaphor/ analogy				•	•
Favourable social & economic conditions					•

In this second category of human ancestors, which includes a variety of species (e.g., *H. ergaster*, *H. erectus*, *H. heidelbergensis*), Mithen (2002) indicates the likelihood of further developments in observational and hypothesis-testing abilities, together with the emergence of a capacity for causal inference. In support of these speculations, Mithen points to evidence of increased sociality, the development of more complex technology, (e.g., emergence of the bifacial technique), and big game hunting (necessarily a co-operative venture), as signs of a significant shift in the mind-reading abilities of early humans compared to those of the early hominines. Moreover, he remarks that such mind-reading skills are intimately tied to an appreciation of basic causal principles and speculates that a concern with causation may have initially emerged in the context of folk psychological reasoning, and only later been extended to natural world phenomena.

Turning to the cognitive foundations of science present in the Neanderthals (*H. neanderthalensis*), Mithen (2002) argues that evidence of hunting, sophisticated tool making techniques (including the Levallois method), and social behaviour (supported by a language facility), can be seen to reinforce claims for the early emergence of basic scientific abilities in human evolution. In addition, he cautiously speculates that one known set of Neanderthal artifacts may constitute an early example of a recording device, thereby indicating that tools were being used by the Neanderthals to extend cognition/perception. However, Mithen also argues that the 'cognitive profile' of Neanderthals emerging from the archaeological record is remarkably static over time, to the extent that "... Neanderthals living at 50,000 years ago appear to have no greater store of knowledge or understanding of the world than those living at 250,000 years ago" (Mithen, 2002, p.32). In Mithen's view, this lack of "directional change" in knowledge, together with the absence of any compelling evidence for art, symbolism, or ritual, suggests that additional key elements of scientific thought are likely to be absent from this group.

In contrast to the static profile of the Neanderthals, Mithen (2002) argues that the appearance of the first anatomically modern humans (*H. sapiens sapiens*) marks the beginning of the period in which all the remaining cognitive foundations of science can be seen to emerge. In particular, he points to dramatic developments at c.50,000 years ago in tool technology, the use of organic substances, body adornment and art, as evidence for a fundamental shift in intellectual abilities relevant to the eventual development of science.⁴ First, he suggests that the archaeological record for this period offers compelling evidence that tools were being used to extend cognition and thereby facilitate problem solving, the most obvious example being incised bones and stones, which are thought to have functioned as rudimentary recording devices or 'external memory aids'. Second, Mithen highlights examples of images from Upper Palaeolithic art that combine features of humans and animals, and argues that they indicate the capacity for analogical reasoning and the use of metaphor. Third, he

⁴ Mithen's specific proposal, developed in detail elsewhere (Mithen, 1996b), is that these developments reflect a major cognitive transition from a "specialized mentality" to a new form of "cognitively fluid mentality", which is a necessary precursor to creative thought and can be seen to underlie the emergence of art, religion, and science.

points to the development of hunting and gathering tools and strategies designed to match prevailing environmental conditions, as well as the convergence on a microlithic hunting technology in many parts of the world, as evidence for substantial accumulation of knowledge over time.

Finally, Mithen (2002) concludes that although all the key elements of scientific thought were in place in late Pleistocene hunter-gatherer societies, the development of science as a distinct domain of inquiry was ultimately dependent on a specific set of social and economic conditions which emerged much later in human history. With this in mind, he turns to the early farming settlements in the Near East to consider the role of the invention of agriculture, citing the emergence of a burgeoning craft culture, and the development of substantial bodies of new knowledge in building, agriculture, and textiles, as evidence of the contribution of a farming economy to the eventual appearance of science.

In contrast to the evolutionary story presented by proponents of the theory theory, Mithen's (2002) account of the cognitive origins of science can be seen to offer a far more comprehensive solution to the puzzle of what makes science possible that succeeds in overcoming the limitations of Gopnik and Meltzoff's 'just-so' account. While the theory theory suggests that our core scientific abilities emerged as a by-product of childhood learning, Mithen rejects attempts to pinpoint a particular time, or even a particular species, as holding the key to our capacity for science. Rather, his proposal is that scientific reasoning constitutes an emergent phenomenon, embodying multiple elements that were subject to different selection pressures and evolved independently of one another over the course of human evolution. As he states, "[t]he human mind is a product of a long evolutionary history" (Mithen, 2002, p.40). Recognizing this fact indicates that it is only by undertaking a detailed investigation of this history in its entirety, examining the fossil and archaeological evidence for clues to the thought and behaviour patterns of our ancestors and relatives, that we will be in a position to gain some purchase on the cognitive origins of science (Mithen, 2002).

6.3 Abduction, tracking, and science

In this chapter, I have examined Gopnik and Meltzoff's (1997) proposal that our capacity for science rests on theory formation capabilities that evolved to facilitate early childhood learning. I began by tracing the development of this evolutionary story across a series of successive formulations of the theory theory, highlighting both the development of cognitive connections between science and childhood, and the claim that science's success is ultimately dependent on these ties with childhood cognition. Following this presentation, I identified a number of criticisms that have been directed at this account, and argued that its 'just-so' status, dependence on the veridicality of human cognition, and inability to account for science's historical development, all serve to undermine Gopnik and Meltzoff's proposed solution to the evolutionary puzzle of our capacity for science.

Having discounted the theory theory's story, I subsequently outlined an alternative account of the cognitive origins of science put forward by Mithen (2002), and backed by archaeological evidence, which suggests that our core scientific reasoning abilities are the emergent outcome of a piecemeal evolutionary history. Sketching the outline of this history from the earliest traces of humanity to the appearance of the first farming communities, I concluded that Mithen's research programme represents an informative approach to the issue of what makes science possible and, in contrast to the theory theory's just-so story, offers a useful framework for future investigations of the cognitive basis of science.

Finally, given the focus of the current work on *abductive* methods of science, this chapter would not be complete without some preliminary speculations about the possible origins of abduction in human thought. In this respect, a recent account of the "roots of scientific reasoning" that examines the degree of continuity across reasoning strategies in human hunter-gatherers and scientists offered by Peter Carruthers (Carruthers, 2002b), is suggestive. Moreover, Carruthers' claims can be seen to be directly relevant to the continuity claims defended in this thesis, because they focus on the species of abductive inference involved in the generation and evaluation of theories that comprise the cornerstones of Haig's (2002) abductive theory of scientific method.

Briefly, Carruthers (2002b) reviews a selection of anthropological data relating to the extended processes of reasoning involved in tracking animals employed by contemporary hunter-gatherers, and argues for the existence of fundamental continuities between these inferential practices and the knowledge seeking strategies of scientists. He begins by identifying the critical elements of scientific reasoning with the processes of existential abduction and inference to the best explanation that were detailed in Chapter 3. Then, working predominantly from a detailed account of tracking and its links to science compiled by Liebenberg (1990),⁵ Carruthers demonstrates that the sophisticated speculative tracking engaged in by hunters in pre-scientific cultures relies on just these forms of abductive inference. In particular, he highlights the need for hunters tracking an animal to construct rudimentary explanatory hypotheses about the likely behaviour of the animal, by reasoning back from traces of footprints, disturbed vegetation, etc. to the causal mechanisms underlying these signs. Moreover, he remarks that once generated, these hypotheses will typically be subjected to a process of intense comparative evaluation, in which hunters draw on a range of empirical and super-empirical criteria in order to establish which of the available hypotheses provides the best explanation of the spoor evidence.

Having endorsed a continuity position on the relations between science and commonsense, Carruthers (2002b) proceeds to outline his preferred version of the emergence of these scientific abilities in human development, and speculates that the appearance of “sophisticated cross-modular abductive reasoning” in school age

⁵ In his book “The art of tracking: the origin of science”, Liebenberg (1990) relates the reasoning involved in tracking practiced by contemporary hunter-gatherers of the Kalahari to a hypothetico-deductive (rather than abductive) model of science, which he then contrasts with an inductive account. A key feature of his argument is that the inferential requirements of sophisticated speculative tracking (as opposed to simple tracking) require the hunter to go beyond enumerative induction to generate a creative hypothesis that coherently explains a range of tracks and signs in terms of underlying causes. However, given that it is this creative generation of a hypothesis or theory and its subsequent evaluation in terms of its ability to provide the best explanation of the spoor evidence that Liebenberg wishes to emphasize in his parallels with scientific practice, I would suggest that the abductive account of scientific method endorsed in this thesis provides a more fitting source model for his claims. In this respect, it is also notable that Carruthers interprets Liebenberg’s findings of hunter-gatherer tracking within an abductive methods framework.

children may be reliant on language as a critical precursor.⁶ Finally, Carruthers concludes with some speculations about the origins of abductive inference in human evolution. In line with Liebenberg's (1990) proposals that the cognitive origins of science reside in the ability to interpret tracks and signs, which in turn may have evolved because of their important role in hunting, Carruthers suggests that one plausible scenario is to tie the adaptive function of abduction to these subsistence activities. In which case, as Carruthers points out, sexual selection forces would have been important in establishing some of the key inferential foundations of science (Carruthers, 2002b).

In summary, while not professing to offer an evolutionary explanation of our capacity for abduction, Carruthers can be seen to offer some interesting speculations about the roots of scientific inference that are relevant to issues at the heart of the child-as-scientist debate. Firstly, his argument for comparable scientific abilities across scientific and pre-scientific cultures endorses the fundamental assumption of continuities between science and commonsense on which the child-as-scientist analogy depends. Secondly, by focusing on the role of abductive methods in science, Carruthers' continuity account provides some measure of support for the specific parallels between scientific and everyday reasoning that are developed and defended in this thesis. Finally, by highlighting the role of existential abduction and inference to the best explanation in contemporary hunter-gatherer tracking, the work of both Carruthers (2002b) and Liebenberg (1990) can be seen to provide a potentially informative platform for future investigations of the origins of abductive science. Working from this platform, researchers can explore the initial speculation put forward by Liebenberg (1990) that "selection for an ability to interpret tracks and signs may have played a significant role in the evolution of the scientific intellect" (Liebenberg, 1990, p.4).

⁶ Given abduction is not dependent on language, a more plausible account of its development may, as indicated in Chapter 4, draw on the relationship between representational redescription and creative theory construction identified by Karmiloff-Smith's (1992) model. Interestingly, this model of the emergence of cognitive flexibility in ontogenetic development has also proven useful for researchers grappling with the problem of how such flexibility may have emerged in human minds over the course of evolutionary history (e.g., see Mithen, 1996a, 1996b; Browne, 1996).

Chapter 7

An Abductive-Methods Perspective on the Child-as-Scientist Debate

I began this thesis by asking how we can best characterize the development of knowledge in childhood. I suggested that, increasingly, researchers' attempts to answer this question are guided by analogy to the development of knowledge in science. As Keil, Levin, Richman and Gutheil (1999) remark, "[i]t has become increasingly tempting in recent years to consider children as intuitive theorists or little scientists" (Keil et al., 1999, p.285). This comment reflects the fact that an extensive body of research now exists that subscribes to the idea that children parallel scientists in their attempts to explain and predict phenomena as they acquire knowledge of the world. Foremost among proponents of this view are Alison Gopnik and Andrew Meltzoff (1997), whose formulation of the theory theory takes the cognitive processes subserving theory change in science and childhood cognitive development to be essentially the same. The aim of this thesis has been to critically examine the theory theory and to draft an alternative version of the analogy in which the construction of meaningful relations between children and scientists is achieved by adopting an abductive-methods perspective on the debate. In this concluding chapter, therefore, I first provide a summary of the main arguments, I then reinforce proposals for the utility of a methodological perspective on child-scientist parallels by undertaking a detailed comparative evaluation the theory theory and the abductive-methods account, and finally close with some speculations regarding future directions.

7.1 Looking back

In order to undertake a critical analysis of the utility of child-scientist parallels, my preparatory work in the Introduction comprised a brief review of the recent history of the child-as-scientist analogy in cognitive developmental research. Beginning with Henry Wellman's application of the scientific analogy to the content and structure of children's everyday knowledge, I highlighted a number of ways in which mappings have been constructed between children and scientists, and sketched some of the criticisms that have been levelled at the analogy by both psychologists and philosophers of science. Having situated the theory theory within its historical context, I then sought to lay the methodological foundations for the model development process to follow by investigating the role of analogy in science. I first identified the features that constitute a productive scientific analogy in general terms, and then outlined a specific theory of analogical reasoning developed by Holyoak and Thagard (1995), which characterizes the mapping process in terms of the simultaneous satisfaction of multiple competing constraints. I suggested that this multiconstraint theory serves to provide some broad directives for my proposed reformulation of child-scientist parallels and, in addition, offers a useful template for undertaking a comparative evaluation of the theory theory and my alternative methods-centred account.

With these methodological provisions for constructing a productive scientific analogy in place, Chapter 1 undertook a detailed examination of Gopnik and Meltzoff's (1997) formulation of the theory theory. I began by outlining the main tenets of the theoretical framework and then focused on the four principal strategies adopted by the authors in a bid to develop the theory theory beyond its current heuristic status. Following this presentation, I turned to consider some of the criticisms that have been directed towards Gopnik and Meltzoff's characterization of child-scientist parallels, concerning both their claims for identity between scientific and childhood cognition, and their extension of the theory theory down to infancy. These criticisms prefigured my proposal that the major limitation of the theory theory account lies in its continued reliance on an inappropriate source model. Specifically, I argued that Gopnik and Meltzoff's account of child-scientist parallels places undue weight on Kuhnian ideas

about revolutionary conceptual change and, as a result, is unable to illuminate the mechanisms responsible for cognitive development in childhood.

Following this critique of the theory theory, Chapter 2 initiated the development of an alternative formulation of child-scientist parallels, by looking to select a more appropriate source model for the analogy. Drawing on the evolutionary naturalist insights of Hooker (1987, 1989), I first outlined the general case for a methods-centred view of science. Then I showed how the expressed purpose of Gopnik and Meltzoff's (1997) theory theory, namely to utilize theory-evidence relations in order to *explain* conceptual and linguistic developments, specifically requires a methods-oriented approach. Having identified the potential utility of a methodological source model for the analogy, I turned to consider possibilities offered by the two major orthodox theories of scientific method. Evaluating each theory in terms of its prescriptions for knowledge generation, I argued that neither an inductive account with its unfulfilled promise of a 'truth-producing algorithm', nor the received hypothetico-deductive view with its emphasis on the sufficiency of empirical testing, has the capacity to provide an adequate account of scientific inquiry. This being the case, I advocated a move beyond orthodox accounts to more recent developments in scientific methodology, in an effort to secure an appropriate source model for my reformulation of child-scientist parallels.

In Chapter 3, I detailed my chosen source model, a comprehensive abductive theory of scientific method put forward by Haig (2002). By way of introducing the abductive theory, I first identified a renewed philosophical interest in the empirical base of science and the role of experimental practice in producing facts worthy of scientific investigation. Following this review of the new experimentalist movement, I highlighted a related body of work by Bogen and Woodward (1988, 1992) that clarifies the nature of empirical facts. These authors demonstrate that it is phenomena, rather than observable data, which are the appropriate focus of scientific explanation and prediction. Due attention to phenomena detection in science in turn suggests the need for a corresponding focus on the explanation of phenomena and the codification of the inferential moves from descriptive claims about empirical phenomena to explanatory theory. In line with this need, I pointed to a growing recognition of abduction as an important species of scientific inference and the

concomitant development of computationally rigorous and operationally specific guidelines for explanatory reasoning. Against this backdrop of contemporary research, I presented Haig's abductive framework, emphasizing its focus on both the detection of empirical phenomena and their subsequent theoretical explanation via an abductive reasoning process that incorporates generation, development, and appraisal dimensions. I concluded that the abductive theory of scientific method offers an informative general perspective on scientific inquiry. Moreover, by doing so, it recommends itself as an appropriate source model with which to reformulate the child-as-scientist analogy.

In order to map relations from the abductive-methods source model to the target of children's knowledge development effectively, Chapter 4 began by redefining the role of the child-as-scientist analogy within cognitive developmental research. I argued that Gopnik and Meltzoff's (1997) attempt to promote the theory theory as a successor to Piaget's constructivist framework does not facilitate the development of system mappings between children and scientists. I then considered as a more promising framework theory for developmental inquiry an alternative interactionist account of development put forward by Elman et al. (1996). Within this framework, I defined a narrower role for the child-as-scientist analogy by reviewing microgenetic analyses of children's problem solving strategies undertaken by Karmiloff-Smith (1984, 1988, 1992) that provide persuasive evidence for the existence of creative theory construction in childhood. I concluded that Karmiloff-Smith's specific findings of data-to-theory moves in children's problem solving which accord with, and are further illuminated by, the abductive account of scientific method, indicate the utility of this reformulation of the analogy for investigating cognitive change.

Having established the initial plausibility of my abductive-methods perspective on the child-as-scientist debate, Chapter 5 extended the application of the model beyond a concern with theory generation to examine the processes by which children evaluate the quality of everyday explanations. With this aim in mind, I reviewed a range of evidence for the claim that school age children are sensitive to criteria for evaluating theory goodness that are commonly identified with theory evaluation practices in science. I then linked these findings to the computational model of theory evaluation advanced by Thagard (1989, 1992), which combines these criteria in inferences to the

best explanation understood in terms of explanatory coherence. Drawing on Samarapungavan's (1992) investigation of children's scientific reasoning strategies, I argued that it is plausible to suggest children can appropriately use explanatory coherence considerations as a basis for theory selection. I concluded that an abductive-methods approach to theory evaluation, which incorporates a precise means of formulating the notion of inference to the best explanation, is capable of illuminating significant parallels between the evaluative strategies deployed in scientific and everyday thought.

Finally, having demonstrated the utility of a methodological reformulation of the child-as-scientist analogy, Chapter 6 turned to speculate on the origins of our theory building abilities in light of Gopnik and Meltzoff's (1997) invocation of an evolutionary warrant for the theory theory. I began by tracing the emergence of an adaptationist explanation of child-scientist parallels through a series of progressive formulations of the theory theory by Gopnik and her colleagues. Following my reconstruction of the theory theory's evolutionary story, I considered a number of criticisms that jointly serve to undermine the theory theory's claim to have identified the cognitive basis of science. Then, having rejected Gopnik and Meltzoff's evolutionary speculations, I reviewed an alternative account of the emergence of scientific cognition in human evolution by Mithen (2002) and concluded that it offers an informative framework for investigating our human capacity for science. Finally, I briefly considered some speculations regarding a possible connection between scientific reasoning and tracking by Liebenberg (1990) and Carruthers (2002b) that hold particular relevance for the methodological parallels drawn in this thesis, and suggested that these accounts may offer a useful entry point for future investigations of the origins of abductive science.

7.2 Comparing analogies: The theory theory versus the abductive-methods account

In her 1996 formulation of the child-as-scientist analogy, Gopnik (1996b) states that in order to move the analogy forward, we need to be more explicit about the proposed parallels between the development of knowledge in science and the development of knowledge in childhood. Likewise, Gopnik and Meltzoff's expressed aim is to present the theory theory in as much detail, and with as much precision as possible, thereby forcing researchers to reconsider its status as merely a "vague metaphor" (Gopnik & Meltzoff, 1997). Yet despite recognizing the need for clarity, Chapter 1's review of recent commentaries on the theory theory highlighted persisting ambiguities, including the strength of the relations being argued for (e.g., Giere, 1996; Stich & Nichols, 1998), and the actual location of the proposed parallels between scientific development and children's cognitive development (Downes, 1999; Bishop & Downes, 2002).

As part of my own preparatory work undertaken in the Introduction, I suggested that attempts to achieve a clearer account of child-scientist parallels would benefit from the store of recent cognitive science research on the nature of analogical reasoning. In particular, I reviewed the multiconstraint theory of analogy put forward by Holyoak and Thagard (1995), and presented a template for undertaking a comparative examination of the theory theory and my alternative abductive-methods account drawn from their comprehensive treatment of analogical thinking. Now being in a position to apply this template, we can draw on the multiconstraint theory's taxonomy of attribute, relational, and system mappings to construct representations of the two analogies, and use the multiple constraints of semantic similarity, structural consistency, and relevance-to-purpose to evaluate the relative coherence of each account.

7.2.1 The theory theory

Table 7.1 offers a representation of Gopnik and Meltzoff's (1997) theory theory using Holyoak and Thagard's tripartite distinction between attribute, relational, and system mappings (Holyoak & Thagard, 1995; see also Shelley, 1999a, 1999b).

Table 7.1 A representation of the theory theory using the multiconstraint theory's taxonomy of attribute, relational, and system mappings.

	SOURCE: SCIENCE	TARGET: CHILDHOOD
Attribute	scientists scientists'-theories scientists'-cognitive-development scientific-theories theoretical-terms theory-change semantic-change scientific-development	children children's-mental-models children's-cognitive-development children's-mental-models children's-early-words conceptual-change semantic-change children's-cognitive-development
Relational	(comprise (theories, coherent-structure) comprise-1) (have (theories, interpretative-function) have-1) (have (theories, predictive-function) have-3) (change (theories) change-1)	(comprise (mental-models, coherent-structure) comprise-2) (have (mental-models, interpretative-function) have-2) (have (mental-models, predictive-function) have-4) (change (mental-models) change-2)
System	(cause (evidence, change-1) cause-1) (entails (change-1, semantic-change) entails-1) (explains (cause-1, scientific-development) explains-1)	(cause (evidence, change-2) cause-2) (entails (change-2, semantic-change) entails-2) (explains (cause-2, cognitive-development) explains-2)

This taxonomy serves to emphasize the depth of the parallels being argued for, and hence is particularly useful for comparing competing analogies. More specifically, each row in the table depicts a particular mapping between the source (science) and the target (childhood). The mappings identified in the top third of the table are *attribute mappings*, where attributes of the source and the target are placed in correspondence with each other on the basis of their perceptual and/or semantic similarity. The middle third of the table depicts more complex *relational mappings*, in which it is the relation between attributes that is mapped from the source analog to the target. Finally, the bottom third of the table contains *system mappings* – the most complex and abstract form of mappings identified by the multiconstraint theory. System mappings constitute mappings of higher order relations based on causal concepts, including ‘cause’, ‘explain’, ‘entail’, ‘facilitate’, etc. and involve whole sets of interconnected relations being mapped from source to target.

The first thing we notice about this representation of the theory theory is that a large number of the proposed correspondences are located at the initial level of attribute mappings. So for example, *scientists* in the source domain maps to *children* in the target domain, and *scientists' theories* maps to *children's mental models*. Moreover, two distinct groupings of attribute mappings can be identified: those that map knowledge structures and processes of scientists as individual cognitive agents to children's knowledge structures and processes, and those that map attributes of science and scientific change to the pattern of knowledge development in childhood. At the next level down, the more complex relational mappings identified are concerned with the content of scientific knowledge and, by analogy, children's knowledge. In particular, these relational mappings identify relations between theories and their characteristic structural, functional, and dynamic properties that are then mapped to children's mental models. So for example, scientific theories comprise coherent structures (*comprise-1*), which have interpretative and predictive functions (*have-1* and *have-3*), and undergo change (*change-1*). Similarly, children's mental models are argued to be coherent (*comprise-2*), to provide children with interpretations and predictions about phenomena (*have-2* and *have-4*), and be defeasible (*change-2*). Finally, the system mappings identified represent the higher order relations invoked by the theory theory to support claims of deep similarities between the processes of knowledge development in science and childhood. As Shelley (1999b) remarks, "... system mappings indicate not just *that* two conceptions are similar, but *why* they are similar (1999b, p.149, italics in original). In the case of the theory theory, the causal-explanatory concepts mapped from source to target can be seen to draw on a broadly Kuhnian analysis of scientific development. The relations being mapped suggest that the accumulation of counter-evidence is primarily responsible for theory change (*cause-1*) and this interplay between theories and evidence, which also entails semantic changes (*entails-1*), accounts for scientific development (*explains-1*). Similarly, Gopnik and Meltzoff (1997) argue that evidence causes children's mental models to change (*cause-2*), which entails associated linguistic developments (*entails-2*), and provides an explanation for the majority of cognitive development in childhood (*explains-2*).

Working from this multiconstraint representation of the theory theory, then, provides an initial indication of the complexity of the parallels constructed by Gopnik and

Meltzoff (1997). At a general level, it indicates that many of the mappings involve shared lists of attributes rather than the more complex causal relations that obtain in system mappings. More specifically, this representation reveals that proposals for substantive *cognitive* parallels between scientists and children, which are central to Gopnik and Meltzoff's (1997) defence, fail to demonstrate any degree of analogical depth. In particular, the attribute mappings constructed between scientists' (as opposed to 'science') and children's knowledge acquisition processes, are not supported by any system mappings that identify the relevant set of causal relations to be transferred from source to target. Rather, the system mappings given in Table 7.1 draw their warrant from Kuhn's analysis of the social/institutional mechanisms involved in revolutionary scientific change (e.g., Kuhn, 1962). The resulting absence of system mappings licensed by a cognitive-oriented theory of scientific inquiry means that the theory theory's construal of child-*scientist* parallels lacks analogical complexity.

According to a multiconstraint analysis, the goodness or coherence of an analogy is determined by applying the standards of similarity, structure, and purpose. More specifically, Holyoak and Thagard (1995) argue that a productive scientific analogy promotes system mappings between source and target by simultaneously satisfying the following constraints:

- i) Semantic Similarity – the degree of perceptual and/or semantic similarity between mapped elements of the source and the target;
- ii) Structural Consistency – the degree to which the analogy constitutes an isomorphism;
- iii) Relevance-to-Purpose – the degree to which the analogy provides a solution to the problem that prompted the construction of the analogy in the first place (Holyoak & Thagard, 1995).

How does the theory theory fare when judged against this set of evaluative criteria? Taking the constraint of similarity first, Gopnik and Meltzoff (1997) can be seen to rely on natural correspondences between components of the source and the target that exploit perceptual and semantic similarities holding between scientists and children. So for example, scientists' conceptual structures and children's conceptual structures

are paired on the basis of object similarity, as are scientist's cognitive processes and children's cognitive processes. In addition, the source and target are seen to be semantically similar: in both cases, theoretical structures are implicated in knowledge acquisition, which in turn is seen to involve a process of theory change. Moreover, Table 7.1 shows that Gopnik and Meltzoff's (1997) relational and system mappings involve identical predicates, thereby indicating that the analogy between the development of knowledge in science and the development of knowledge in childhood is a strong and reasonable one.

Countering this conclusion, however, are the criticisms of the theory theory reviewed in earlier chapters, many of which can be seen to relate to the constraint of similarity. In particular, researchers such as Bishop and Downes (2002), Downes (1999), Gellatly (1997), and Russell (1992), all question in various ways the validity of the proposal that revolutionary theory change in science and cognitive development in childhood are similar or even comparable processes given that they rely on fundamentally different sets of causal relations for their operation. In Chapter 1, I argued that Gopnik and Meltzoff's (1997) attempt to counter these concerns by giving science a cognitive characterization and attempting to promote parallels between scientists and children as individual cognitive agents. Based on the representation of the theory theory given in Table 7.1, however, this criticism appears to be upheld. That is, while claiming to have identified substantive cognitive correspondences between children and scientists, the majority of the mappings constructed by Gopnik and Meltzoff (1997) can be seen to involve non-cognitive comparisons between abstract features of science and scientific change and children's knowledge development. Moreover, the theory theory's system mappings that indicate the reasons why the source and target are similar, attach identical predicates to an institutional process of knowledge change on the one hand, and an individual's knowledge acquisition on the other, thereby violating constraints of similarity at the level of causal relations.

In addition to a failure to satisfy similarity constraints at the level of system mappings, the theory theory has also been criticized on grounds of structural inconsistencies. In particular, work by Bishop and Downes (2002) and Downes (1999) suggests that the theory theory is "multiply ambiguous", promoting a range of possible relations

between the source analog and the target that compete with one another and create tensions within the analogy. According to the multiconstraint theory, the constraint of structural consistency is satisfied in an analogy when two conditions are met: 1) each predicate in the source analog is aligned with a unique predicate in the target; and 2) when two predicates are mapped, their corresponding arguments are also mapped from source to target. Table 7.1 reflects the concerns raised by Bishop and Downes, by depicting attribute mappings that violate the first of these conditions. Specifically, Gopnik and Meltzoff (1997) can be seen to align *children's cognitive development* with both *scientists' cognitive development* and *scientific development*, and *children's mental models* with *scientists' theories* and *scientific theories*, thereby violating the one-to-one mapping constraint that impacts on the theory theory's overall structural integrity (Holyoak & Thagard, 1995; see also Gentner, 1983).

Finally, my own criticisms of the theory theory can be seen to relate to the third constraint on analogical reasoning identified by the multiconstraint theory, which concerns the purpose of the analogy. Discussing this constraint, Holyoak and Thagard (1995) argue that it is important to examine the elements of the source analog that are relevant to the user's goal of solving a problem in the target domain. For example, in explanation-driven uses of analogy, the relevant elements are the set of causal relationships operating in the source that can point to potential causes of the target behaviour under study (Holyoak & Thagard, 1995). In Chapters 1 and 2, I suggested that Gopnik and Meltzoff's (1997) goal in using the analogy is to develop a replacement framework theory for cognitive development that provides an explanatory account of the cognitive mechanisms subserving conceptual and linguistic developments in children. From the representation given in Table 7.1, however, the theory theory version of the analogy fails to meet this goal, due to its continued reliance, depicted in its system mappings, on a Kuhnian account of scientific development. As I have repeatedly argued, the set of causal relationships operating in this account offers no insights into individual scientists' cognition and, accordingly, is unable to contribute to productive reasoning about the mechanisms underpinning children's knowledge acquisition processes.

7.2.2 The abductive-methods account

Contrasting with the representation of the theory theory given in Table 7.1, Table 7.2 details the particular pattern of attribute, relational, and system mappings underpinning the abductive-methods version of the child-as-scientist analogy.

Table 7.2 A representation of the abductive-methods account using the multiconstraint theory's taxonomy of attribute, relational, and system mappings.

	SOURCE	TARGET
Attribute	scientists scientists'-theories scientists'-methods	children children's-mental-models children's-methods
Relational	(comprise (theories, coherent-structure) comprise-1) (have (theories, explanatory-function) have-1) (construct (scientists, theories) construct-1) (employ (scientists, existential-abduction) employ-1) (apply (scientists, EC-criteria) apply-1)	(comprise (mental-models, coherent-structure) comprise-2) (have (mental-models, explanatory-function) have-2) (construct (children, models) construct-2) (employ (children, existential-abduction) employ-2) (apply (children, EC-criteria) apply-2)
System	(in-order-that (construct-1, explain) in-order-that-1) (enables (employ-1, theory-generation) enables-1) (facilitates (apply-1, theory-evaluation) facilitates-1) (explains (enables-1, construct-1) explains-1) (explains (facilitates-1, construct-1) explains-3)	(in-order-that (construct-2, explain) in-order-that-2) (enables (employ-2, model-generation) enables-2) (facilitates (apply-2, model-evaluation) facilitates-2) (explains (enables-2, construct-2) explains-2) (explains (facilitates-2, construct-2) explains-4)

The first point of difference concerns the relative distribution of mappings across the three categories, with the majority of correspondences between the source and the target on the abductive-methods account being located at the levels of relational and system mappings. Further, the mappings identified are more tightly focused on scientists' cognitive constructions and their methods of inquiry, with the attribute mappings between *scientists' theories* and *children's mental models*, and between *scientists' methods* and *children's methods* being supported by deeper relational mappings that identify the specific inferential strategies involved. In particular, the relational mappings shown can be seen to go beyond the proposal that both scientists and children construct theories to include additional correspondences indicating that

scientists employ existential abduction (*employ-1*) and that they apply explanatory coherence criteria (*apply-1*). Similarly, children are argued to rely on the same abductive principles in their problem solving attempts, employing existential abduction (*employ-2*) and applying criteria relevant to explanatory coherence considerations (*apply-2*).

Finally, the system mappings promoted by the abductive-methods account offer additional support for the attribute and relational mappings identified between scientists and children, by giving the reasons for the lower level relations between the source and the target. These include the motivations to engage in theory building activities; that is, scientists and, by analogy, children are seen to construct their representational structures in order to gain explanatory understanding (*in-order-that*). Further, the capacity for abductive reasoning is directly related to the ability to effectively generate and evaluate claims about theoretical entities. Scientists' employment of existential abduction enables them to generate explanatory theories (*enables-1*), and their application of explanatory coherence criteria facilitates the effective evaluation of their theories (*facilitates-1*), which together contribute to an explanation of the theory formation process (*explains-1* and *explains-3*). Similarly, children's ability to use existential abduction and apply explanatory coherence criteria allow them to generate and evaluate explanatory models (*enables-2* and *facilitates-2*), which combine to offer the beginnings of an explanation of how children develop knowledge of the world (*explains-2* and *explains-4*).

Compared with the multiconstraint representation of the theory theory, the abductive-methods account depicted in Table 7.2 demonstrates a greater level of mapping complexity in its comparisons of scientists and children. To begin with, the majority of the mappings depicted are relational and system mappings as opposed to more superficial mappings between basic attributes of the source and target. In addition, the abductive-methods account resolves the tension between theory theory comparisons based on the scientist as individual cognitive agent and those based on an abstract account of scientific change, by focusing exclusively on methodological parallels between scientists and children informed by an abductive theory of scientific method. As a result, the methods-centred mappings constructed between scientists and children at both attribute and relational levels are supported and further deepened

by complex system mappings that apply causal-explanatory concepts drawn from the abductive theory to make sense of the theory construction efforts of children.

The proposal that the abductive-methods version of child-scientist parallels demonstrates a greater degree of mapping complexity than the theory theory gains further credence when judged against the standards of the multiconstraint theory. Firstly, regarding the constraint of similarity, this alternative abductive-methods account can be seen to overcome many of the concerns raised by theory theory critics relating to violations of this constraint. Like the theory theory, the abductive-methods account is initially guided by natural correspondences between components of the source and the target that draw on perceptual and semantic similarities holding between scientists and children. However, unlike the theory theory, this alternative account is able to capitalize on these correspondences by adopting a methodological source model that searches for a deeper basis to these similarities. In addition, this source model allows the abductive-methods account to avoid many of the fundamental dissimilarities between source and target facing theory theory proponents who are attempting to match the process of knowledge development in children to an abstract, disembodied account of scientific theory change. Furthermore, the use of identical predicates by the abductive-methods account, indicating that the source and target are semantically very similar, is supported by the range of evidence reviewed in Chapters 4 and 5 that points to compelling similarities between the inferential strategies actually employed by scientists and children in their knowledge seeking endeavours.

Secondly, by selecting a methodological source model that gives precedence to the view of scientists as cognitive agents, the abductive-methods account avoids the structural inconsistencies identified in the theory theory analogy. In Gopnik and Meltzoff's (1997) account, tendencies to violate one-to-one mappings can be seen to stem from the inability of a Kuhnian source model to support the effective construction of cognitive mappings between scientists and children. As a result, children's representational structures and their knowledge acquisition processes are not only mapped to scientific theories and scientific development, but also to scientists' theories and scientists' cognitive development, thereby weakening the structural integrity of the analogy. In contrast, the abductive-methods account, with

its focus on the methods and inferential strategies employed by individual scientists, eliminates the need for such two-to-one mappings. Specifically, Table 7.2 shows that each element in the source is mapped to exactly one element in the target, thereby ensuring that the abductive-methods analogy is structurally consistent.

Finally, this alternative abductive-methods version of the child-as-scientist analogy fulfills the purpose for which it was designed, by supporting my proposal of substantive methodological correspondences between scientific and everyday cognition. In Chapter 4, I made it clear that I did not endorse Gopnik and Meltzoff's (1997) goal of constructing a replacement explanatory framework for cognitive developmental research. Rather, I argued for a revised role for the analogy, in which the focus of investigation is limited to identifying correspondences between the explanatory reasoning strategies employed by scientists and children in their attempts to construct accurate representations of the world. Therefore, in one important sense, the purpose of the abductive-methods analogy is far narrower and more constrained than that of the theory theory, dealing exclusively with abductive principles of inference and attempting to transfer knowledge about the operation of these principles in a scientific context to the everyday context of children's reasoning. However, in another sense this alternative formulation can be seen to have a much broader goal in mind than the goal articulated by Gopnik and Meltzoff (1997). Rather than seeking to identify a 'special' relationship between children and scientists, the abductive-methods account attempts to identify robust commonalities in method that extend across science and commonsense inquiry. From the representation given in Table 7.2, this goal is achieved by the construction of systematic correspondences between patterns of higher order relations that provide an integrated abductive-methods interpretation of creative reasoning in scientific and everyday contexts.

To summarize, by employing a multiconstraint analysis of the theory theory and the alternative methods-centred account, we gain a rich and informative way of representing and evaluating these two analogies. Using Holyoak and Thagard's (1995) taxonomy of attribute, relational, and system mappings to depict each formulation of child-scientist parallels serves to highlight the complexity of the causal connections made by the abductive-methods account relative to the theory theory. Moreover, this indication of greater complexity and coherence is reinforced by the

degree to which the abductive-methods analogy satisfies the constraints of similarity, structure, and purpose. Specifically, based on its ability to construct mappings of similar higher-order relations that evidence a high degree of structural consistency and contribute to its purpose, we can conclude the abductive-methods analogy offers a means of developing relations of greater abstractness, generality, and complexity between scientific and everyday cognition. These relations promise to advance the child-as-scientist debate beyond current theory theory concerns in a number of potentially fruitful ways.

7.3 Future directions: Consolidating an abductive perspective on human reasoning

I have argued that Gopnik and Meltzoff (1997) look to the analogy with science to answer the question of how knowledge develops. In particular, they aim to extend the analogy beyond its widespread application to the content of everyday conceptual knowledge, to investigate the processes by which children form and revise this knowledge over the course of development. According to Gopnik and Meltzoff, children are not just in possession of conceptual frameworks that look a lot like theories, they *are* theoreticians, closely resembling scientists in their attempts to explain and predict the phenomena encountered in their everyday interactions with the world. Having undertaken an extended analysis of the theory theory and its implications for cognitive developmental research, I have argued that the theory theory fails in its bid to offer a coherent and productive account of this knowledge development process. Despite the authors' claims that the theory theory can explain many important psychological phenomena of cognitive development, I have shown that the theory theory's reliance on a Kuhnian source model of scientific change does not facilitate the construction of system mappings between scientists and children. In the absence of any system mappings supporting substantive cognitive parallels, the theory theory is unable to show how insights about the causal relations at work in the source, answer questions about the corresponding set of causal relations in the target.

In addition to analysing the theory theory, however, I have also attempted to formulate a more viable version of the child-as-scientist analogy, based on methodological parallels with the processes of inquiry identified in scientific practice. Focusing on the application of abductive inference strategies not only creates informative connections between the creative reasoning of scientists and children, but also points the way to the formation of a more general category or abductive *schema* that both includes these two examples within its scope and applies to investigations of human reasoning more widely. In their account of the role of analogy in creative thought, Holyoak and Thagard (1995) emphasize the connection between analogy and the formation of schemas, which they characterize as complex concepts "... based on the relational structure common to the target and the source" (Holyoak & Thagard, 1995, p.220). They propose a cyclical relationship whereby the construction of analogies facilitates the formation of new and more abstract schemas, which in turn promote the discovery of more remote analogies, and so on. Likewise, Nersessian (1999) stresses the importance of generic abstraction in analogical modelling, arguing that the process of abstracting what is common to specific instances in a problem context in order to form a general category based on patterns of higher order relations, plays a crucial role in scientific discovery and conceptual change. Relating these proposals to the child-as-scientist debate, the abductive-methods account of the analogy can be seen to draw on insights from scientific methodology, cognitive developmental research, and computational philosophy of science, to abstract what scientists' and children's reasoning strategies have in common in the context of creative theory construction. The result is the identification of a pattern of relations between species of abductive inference and their functional roles in generating and evaluating explanations of empirical phenomena that has potential applications beyond the specific instances examined in this thesis. Speculating on how such a schema could contribute to the consolidation of an abductive perspective on human reasoning more generally, the following lines of inquiry would appear particularly promising:

A cognitive developmental perspective on abduction can complement current philosophical and AI approaches In Chapter 3, I pointed to a recent revival of interest in abductive inference by both philosophers of science and artificial intelligence researchers. Within philosophy of science, creative reasoning strategies, long

marginalized under a deductive logic model of scientific reasoning, have received increased attention in attempts to codify scientific discovery practices (e.g., Curd, 1980; Magnani, Nersessian & Thagard, 1999), and have been accorded a central role in attempts to construct an adequate model of scientific inference more generally (e.g., Haig, 2002; Magnani, 2001; Thagard, 1988; Rozeboom, 1972, 1997). Within the arena of AI, the development and testing of abductive solutions to a range of problems including medical diagnosis, planning, and decision-making, has led researchers such as Josephson and Josephson (1994) to argue that abduction is ubiquitous in human reasoning, and that abductive inference models have potential for wide application across scientific and everyday contexts.

Adding to the results of this “renaissance of research on explanatory reasoning” (Thagard, 2001), the current work has revealed substantive correspondences between scientists and young children in the abductive inferential strategies underpinning creative theory construction. In light of these correspondences, a cognitive developmental approach to the study of abduction could prove a useful source of insights that serves to complement the existing lines of investigation by philosophers and AI researchers. Specifically, detailed empirical investigations of the emergence of abductive inference in human development could contribute to an improved understanding of the necessary precursors to creative reasoning, which in turn could inform the design of computational systems that attempt to model human discovery processes. In addition, a focus on the earliest forms of abductive inference in young children may assist in answering questions of whether or not there are more basic or fundamental kinds of explanatory reasoning and, if so, how these forms relate to the species of abductive inference identified in scientific practice. Such answers could contribute to current attempts to develop taxonomies of abductive inference (e.g., Magnani, 2001; Thagard, 1988), as well as provide a starting point for tracing the developmental trajectory of explanatory reasoning capabilities more generally. Conversely, ongoing work carried out by cognitive developmentalists under the broad heading of causal and explanatory reasoning, could gain a deeper, more theoretically informed focus by being integrated with current philosophy of science and AI work in an interdisciplinary study of abduction.

Investigating the origins of abductive science may clarify Peirce's speculation that we possess an evolved facility for "guessing right" In Chapter 6 I suggested that a comprehensive solution to the puzzle of what makes science possible would, following Mithen's lead, need to look at evidence for the evolution of our scientific abilities contained in the fossil and archaeological records. Following my review of Mithen's account of the emergence of scientific reasoning in human evolution, I briefly sketched some speculations relating to the possible origins of abductive inference that draw on anthropological studies of pre-scientific cultures. In particular, the detailed fieldwork undertaken by Liebenberg (1990) suggests the presence of some fundamental continuities between the reasoning processes involved in the speculative tracking of contemporary hunter-gatherers and the abductive inferential practices at the heart of scientific practice, thereby linking our evolved human capacity for abduction to its role in successful hunting.

These speculations about the possible origins of abduction in human thought would appear to mesh well with Peirce's own ideas about our capacity for abductive inference. In particular, Peirce believed that the human mind demonstrates a peculiar affinity with nature, to the extent that we demonstrate a facility for making correct abductions or 'guessing right' after only a limited number of attempts:

Nature is a far vaster and less clearly arranged repertory of facts than a census report; and if men had not come to it with special aptitudes for guessing right, it may well be doubted whether in the ten or twenty thousand years that they may have existed their greatest mind would have attained the amount of knowledge which is actually possessed by the lowest idiot. (Peirce, 1932, vol.2, p.476)

Moreover, he suggested that an explanation for the existence of this abductive facility is likely to be found in an evolutionary history of the human mind:

There can, I think, be no reasonable doubt that man's mind, having been developed under the influence of the laws of nature, for that reason naturally thinks somewhat after nature's pattern. (Peirce, 1958, vol.7, p.30)

Future investigations of the possible links between abduction and tracking in human evolution that adopt Mithen's evidence-based approach to studying the cognitive origins of science, then, may be able to clarify Peirce's speculation that we possess an evolved facility for "guessing right". In particular, Mithen's own work (Mithen, 1988) has provided direct evidence for the sorts of abductively driven speculative tracking discussed by Carruthers and Liebenberg, in the imagery of Upper Paleolithic cave art. Working from these sorts of records, and with a comprehensive account of abduction in hand, it remains the task of future researchers to determine whether the tentative connections to tracking can provide the foundations of an evolutionary explanation of abductive reasoning and, if so, identify when during the course of human evolution this facility may have emerged.

Increased knowledge of the abductive inferential strategies underpinning creative theory construction has potential applications for science education

Reviewing the literature on the child-as-scientist analogy reveals that it is not only developmental psychologists who have placed undue weight on Kuhn's ideas about revolutionary conceptual change. Science education researchers can also be seen to have relied heavily on the Kuhnian-informed analogy with science. In a seminal paper, Rosalind Driver and Jack Easley (Driver & Easley, 1978) proposed that students' everyday knowledge of natural phenomena is best understood as coherent bodies of intuitive ideas or "alternative frameworks", rather than misunderstandings or immature versions of adult scientific understandings. In addition, these researchers argued that children's cognitive development could be likened to a series of conceptual revolutions, paralleling Kuhn's (1962) description of paradigm shifts in science, and suggested that this Kuhnian-inspired model of radical theory change may have application for science education research and teaching practices (see also Posner, Strike, Hewson & Gertzog, 1982). This work is generally regarded as an important instigator of the constructivist movement in science education that has dominated contemporary approaches to research and teaching of science since the 1980's (Kyle, Osbourne, Leach, Scott & Norris, 1998).

However, while helpful in alerting science educators to the actual content and structure of children's ideas about natural phenomena, the 'conceptual change as

theory change' model has proved less successful in supplying crucial insights about the processes through which new ideas are generated and evaluated. This has led to increased debate about its pedagogical utility for science education.¹ Given the inability of the Kuhnian-inspired model to provide an account of the psychological mechanisms involved in creative theory construction, the alternative methods-centred account of child-scientist parallels formulated in this thesis could offer a useful starting point for researchers wanting to characterize the process of knowledge acquisition in science students. In particular, delineating close connections between species of abductive inference and their roles in generating and evaluating explanations of empirical phenomena in both scientists and children suggests that an abductive-methods framework could take researchers beyond the basic idea that children construct theories to offer some specific guidelines about the inferential strategies and methods involved. Incorporating these abductive inference guidelines into their models of cognitive change could in turn prove useful for researchers tackling the complex problem of how children learn science, and guide thinking about how we can best apply knowledge of children's inquiry methods in formal science instruction.

In conclusion, the current work has attempted to offer an informative methodological perspective on the child-as-scientist debate that overcomes limitations inherent in current approaches. The result is an abductive-methods formulation that integrates insights from scientific methodology, cognitive development, and contemporary philosophy of science to identify substantive parallels between the processes by which children and scientists develop knowledge of the world. The widespread failure by philosophers and psychologists until recently to acknowledge the pervasiveness of abductive reasoning in scientific inquiry and everyday problem solving has been, in my view, to the detriment of our understanding of human creativity in both domains. However, by acknowledging that a codifiable form of creative inference exists, and that it is ubiquitous in scientific and everyday knowledge seeking, we can look to advance knowledge of our human methods of discovery and explanation. In this respect, the most important benefits of the child-as-scientist analogy may be to promote the cross-fertilization of ideas between scientific methodologists and

¹ See for example two recent special issues of *Science & Education*: "Children's Theories and Scientific Theories" (September, 1999) and "Thomas Kuhn and Science Education (January, 2000).

cognitive and developmental psychologists on the nature of creative reasoning. Only by detailed, interdisciplinary investigation of abductive inference practices in science and commonsense will we be in a position to say how knowledge development is possible, whether in our everyday exploits as children or in our intellectual pursuits as professional scientists.

References

- Ackermann, R. (1989). The new experimentalism. *British Journal for the Philosophy of Science*, 40, 185-190.
- Bereiter, C., & Scardamalia, M. (1989). When weak explanations prevail. *Behavioral and Brain Sciences*, 12, 468-469.
- Bishop, M. A., & Downes, S. M. (2002). The theory theory thrice over: The child as scientist, Superscientist or social institution? *Studies in History and Philosophy of Science*, 33, 121-136.
- Bogen, J., & Woodward, J. (1988). Saving the phenomena. *Philosophical Review*, 97, 303-52.
- Bogen, J., & Woodward, J. (1992). Observations, theories and the evolution of the human spirit. *Philosophy of Science*, 59, 590-611.
- Bradie, M. (1986). Assessing evolutionary epistemology. *Biology and Philosophy*, 1, 401-459.
- Brewer, W. F., Chinn, C. A., & Samarapungavan, A. (2000). Explanation in scientists and children. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 279-298). Cambridge, MA: MIT Press.
- Brewer, W. F., & Samarapungavan, A. (1991). Children's theories vs. scientific theories: Differences in reasoning or differences in knowledge? In R. R. Hoffman & D. S. Palermo (Eds.), *Cognition and the symbolic processes: Applied and ecological perspectives* (pp. 209-232). Hillsdale, NJ: Erlbaum.
- Browne, D. (1996). Cognitive versatility. *Minds and Machines*, 6, 507-523.

- Campbell, D. T. (1974). Evolutionary epistemology. In P. A. Schilpp (Ed.), *The philosophy of Karl Popper* (pp 413-463). La Salle, IL: Open Court.
- Campbell, N. R. (1920). *Physics: The elements*. Cambridge: Cambridge University Press.
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge: MIT Press.
- Carey, S. (1988). Conceptual differences between children and adults. *Mind and Language*, 3, 167-181.
- Carey, S. (1992). The origin and evolution of everyday concepts. In R. N. Giere (Ed.), *Minnesota Studies in the Philosophy of Science: Vol. 15. Cognitive models of science* (pp. 89-128). Minneapolis: University of Minnesota Press.
- Carruthers, P. (2002a). Human creativity: Its cognitive basis, its evolution, and its connections with childhood pretence. *British Journal for the Philosophy of Science*, 53, 225-249.
- Carruthers, P. (2002b). The roots of scientific reasoning: Infancy, modularity and the art of tracking. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 73-95). Cambridge: Cambridge University Press.
- Carruthers, P., Stich, S., & Siegal, M. (2002). *The cognitive basis of science*. Cambridge: Cambridge University Press.
- Chalmers, A. F. (1999). *What is this thing called science?* (3rd ed.). St. Lucia, Qld: University of Queensland.
- Chatfield, C. (1985). The initial examination of data. *Journal of the Royal Statistical Society, Series A*, 148, 214-253.
- Chi, M. T. H. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. N. Giere (Ed.), *Minnesota*

- Studies in the Philosophy of Science: Vol. 15. Cognitive models of science* (pp 129-186). Minneapolis: University of Minnesota Press.
- Clark, A., & Karmiloff-Smith, A. (1993). The cognizer's innards: A psychological and philosophical perspective on the development of thought. *Mind and Language*, 8, 487-519.
- Curd, M. V. (1980). The logic of discovery: An analysis of three approaches. In T. Nickles (Ed.), *Scientific discovery, logic, and rationality* (pp. 201-219). Dordrecht: Reidel.
- diSessa, A. (1983). Phenomenology and the evolution of intuition. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 15-33). Hillsdale, NJ: Erlbaum.
- diSessa, A. (1988). Knowledge in pieces. In G. Forman & P. B. Pufall (Eds.), *Constructivism in the computer age* (pp. 49-70). Hillsdale, NJ: Erlbaum.
- diSessa, A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10, 105-225.
- Downes, S. M. (1999). Can scientific development and children's cognitive development be the same process? *Philosophy of Science*, 66, 565-578.
- Driver, R., & Easley, J. (1978). Pupils and paradigms: A review of literature related to concept development in adolescent science students. *Studies in Science Education*, 5, 61-84.
- Duhem, P. (1954). *The aim and structure of physical theory*. Princeton, NJ: Princeton University Press.
- Dunbar, K. (1999). How scientists build models: InVivo science as a window on the scientific mind. In L. Magnani, N. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery* (pp. 85-99). New York: Kluwer Academic /Plenum.

- Dunbar, R. (1996). *The trouble with science*. London: Faber and Faber.
- Efron, B., & Gong, G. (1983). A leisurely look at the bootstrap, the jackknife and cross-validation. *American Statistician*, 37, 36-48.
- Einstein, A. (1934). *The world as I see it*. New York: Covici-Friede.
- Eliasmith, C., & Thagard, P. (1997). Waves, particles, and explanatory coherence. *British Journal for the Philosophy of Science*, 48, 1-19.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
- Faucher, L., Mallon, R., Nazer, D., Nichols, S., Ruby, A., Stich, S., & Weinberg, J. (2002). The baby in the lab-coat: Why child development is not an adequate model for understanding the development of science. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 335-362). Cambridge: Cambridge University Press.
- Feyerabend, P. (1975). *Against method*. London: New Left Books.
- Fine, A. (1996). Science as child's play: Tales from the crib. *Philosophy of Science*, 63, 534-537.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Fodor, J. (1997, May 16). Do we have it in us? The failure of new attempts to explain how the cognitive brain works. *Times Literary Supplement*, pp. 3-4.
- Forbus, K. D., Gentner, D., & Law, K. (1994). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19, 141-205.

- Franklin, A. (1986). *The neglect of experiment*. Cambridge: Cambridge University Press.
- Franklin, A. (1990). *Experiment, right or wrong*. Cambridge: Cambridge University Press.
- Galison, P. (1987). *How experiments end*. Chicago: University of Chicago Press.
- Gellatly, A. (1997). Why the young child has neither a theory of mind nor a theory of anything else. *Human Development*, 40, 32-50.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). Cambridge: Cambridge University Press.
- Giere, R. N. (1988). *Explaining science: A cognitive approach*. Chicago: University of Chicago Press.
- Giere, R. N. (Ed.). (1992). *Cognitive models of science. Minnesota Studies in the Philosophy of Science: Vol. 15*. Minneapolis: University of Minnesota Press.
- Giere, R. N. (1996). The scientist as adult. *Philosophy of Science*, 63, 538-541.
- Glymour, C. (1980). *Theory and evidence*. Princeton: Princeton University Press.
- Gopnik, A. (1984). Conceptual and semantic change in scientists and children: Why there are no semantic universals. *Linguistics*, 20, 163-179.
- Gopnik, A. (1988). Conceptual and semantic development as theory change: The case of object permanence. *Mind and Language*, 3, 197-216.

- Gopnik, A. (1996a). The post-Piaget era. *Psychological Science*, 7, 221-225.
- Gopnik, A. (1996b). The scientist as child. *Philosophy of Science*, 63, 485-514, and Reply to commentators, 552-561.
- Gopnik, A., & Meltzoff, A. N. (1997). *Words, Thoughts, and Theories*. Cambridge, MA: Bradford/MIT Press.
- Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, 7, 145-171.
- Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257-293). New York: Cambridge University Press.
- Goswami, U., & Brown, A. L. (1990a). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, 35, 69-95.
- Goswami, U., & Brown, A. L. (1990b). Higher-order structure and relational reasoning: Contrasting analogical and thematic relations. *Cognition*, 36, 207-226.
- Hacking, I. (1983). *Representing and intervening*. Cambridge: Cambridge University Press.
- Haig, B. D. (1987). Scientific problems and the conduct of research. *Educational Philosophy and Theory*, 19, 22-32.
- Haig, B. D. (1996). Grounded theory as scientific method. *Philosophy of Education 1995: Current issues* (pp. 281-290) Urbana: University of Illinois Press.
- Haig, B. D. (2002). *Towards an abductive theory of scientific method*. Unpublished manuscript, University of Canterbury, Christchurch, New Zealand.

- Haig, B. D., & Durrant, R. (2000). Theory evaluation in evolutionary psychology. *Psychological Inquiry*, 11, 34-38.
- Hanson, N. R. (1958). *Patterns of discovery*. Cambridge: Cambridge University Press.
- Harman, G. (1965). The inference to the best explanation. *The Philosophical Review*, 74, 88-95.
- Harré, R. (1976). The constructive role of models. In L. Collins (Ed.), *The use of models in the social sciences* (pp 16-43). London: Tavistock.
- Harré, R. (1978). Models in science. *Physics Education*, 13, 275-278.
- Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the philosophy of science*. New York: Free Press.
- Hempel, C. G. (1966). *Philosophy of natural science*. Englewood Cliffs, NJ: Prentice Hall.
- Hesse, M. B. (1966). *Models and analogies in science*. Notre Dame, Indiana: University of Notre Dame Press.
- Holyoak, K. J., & Thagard, P. (1989a). Analogical mapping by constraint satisfaction. *Cognitive Science*, 13, 295-355.
- Holyoak, K. J., & Thagard, P. (1989b). A computational model of analogical problem solving. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 242-266). Cambridge: Cambridge University Press.
- Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.
- Hooker, C. A. (1987). *A realistic theory of science*. Albany: State University of New York Press.

- Hooker, C. A. (1989). Evolutionary epistemology and naturalist realism. In K. Hahlweg & C. A. Hooker (Eds.), *Issues in evolutionary epistemology*, (pp. 101-150). Albany: State University of New York Press.
- Hoyningen-Huene, P. (1993). *Reconstructing scientific revolutions*. Chicago: University of Chicago Press.
- Johnson, M. H. (1998). The neural basis of cognitive development. In W. Damon (Series Ed.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 1-49). New York: Wiley.
- Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development*. Oxford: Blackwell.
- Josephson, J. R., & Josephson, S. G. (Eds.). (1994). *Abductive inference*. Cambridge: Cambridge University Press.
- Karmiloff-Smith, A. (1984). Children's problem solving. In M. E. Lamb, A. L. Brown, & B. Rogoff (Eds.), *Advances in developmental psychology: Vol. 3*. (pp. 39-90). Hillsdale, NJ: Erlbaum.
- Karmiloff-Smith, A. (1988). The child is a theoretician not an inductivist. *Mind and Language*, 3, 183-197.
- Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Karmiloff-Smith, A. (1994). Précis of Beyond modularity: A developmental perspective on cognitive science. *Behavioral and Brain Sciences*, 17, 693-745.
- Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, 2, 389-398.

- Karmiloff-Smith, A., & Inhelder, B. (1974). If you want to get ahead, get a theory. *Cognition*, 3, 195-212.
- Keenan, T., Olson, D. R., & Marini, Z. (1998). Working memory and children's developing understanding of mind. *Australian Journal of Psychology*, 50, 76-82.
- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- Keil, F. C. (1998). Cognitive science and the origins of thought and knowledge. In W. Damon (Series Ed.) & R. M. Lerner (Vol. Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development*. (5th ed., pp. 341-413). New York: Wiley.
- Keil, F. C., Levin, D. T., Richman, B. A., & Gutheil, G. (1999). Mechanism and explanation in the development of biological thought: The case of disease. In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 285-319). Cambridge, MA: MIT Press.
- Kemp, S. (1990). *Medieval psychology*. New York: Greenwood Press.
- Kempton, W. (1987). Two theories of home heat control. In D. Holland & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 222-242). Cambridge: Cambridge University Press.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96, 674-689.
- Kuhn, D., Amsel, E., & O'Loughlin, M. (1988). *The development of scientific thinking skills*. San Diego: Academic Press.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Kuhn, T. S. (1977). *The essential tension*. Chicago: University of Chicago Press.
- Kuhn, T. S. (1982). Commensurability, comparability, communicability. *PSA 1982, Vol. 3*. (pp. 669-688). East Lansing, Mich.: Philosophy of Science Association.
- Kyle, W. C., Osbourne, J., Leach, J., Scott, P., & Norris, S. P. (1998). Editor's note. *Journal of Research in Science Teaching*, 35, 1-2.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91-195). Cambridge: Cambridge University Press.
- Laudan, L. (1977). *Progress and its problems: Towards a theory of scientific growth*. Berkeley: University of California Press.
- Leslie, A. (1987). Pretense and representation: The origins of "theory of mind." *Psychological Review*, 94, 412-426.
- Leslie, A. M. (1992). Pretense, autism, and the "Theory of Mind" module. *Current Directions in Psychological Science*, 1, 18-21.
- Levine, A. T. (2000). Which way is up? Thomas S. Kuhn's analogy to conceptual development in childhood. *Science & Education*, 9, 107-122.
- Liebenberg, L. (1990). *The art of tracking: The origin of science*. Cape Town: David Philip.
- Magnani, L. (2001). *Abduction, reason, and science: Processes of discovery and explanation*. New York: Kluwer Academic/Plenum.

- Magnani, L., Nersessian, N., & Thagard, P. (Eds.). (1999). *Model-based reasoning in scientific discovery*. New York: Kluwer Academic/Plenum.
- Mayo, D. G. (1996). *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press.
- McCauley, R. N. (1986). Intertheoretic relations and the future of folk psychology. *Philosophy of Science*, 53, 179-199.
- McCloskey, M., & Kargon, R. (1988). The meaning and use of historical models in the study of intuitive physics. In S. Strauss (Ed.), *Ontogeny, phylogeny, and historical development* (pp. 49-67). Norwood, NJ: Ablex.
- Mill, J. S. (1872). *A system of logic ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation* (8th ed.). London: Longmans.
- Millgram, E., & Thagard, P. (1996). Deliberative coherence. *Synthese*, 108, 63-88.
- Mithen, S. (1988). Looking and learning: Upper Palaeolithic art and information gathering. *World Archaeology*, 19, 297-327.
- Mithen, S. (1996a). The early prehistory of human social behaviour: Issues of archaeological inference and cognitive evolution. *Proceedings of the British Academy*, 88, 145-177.
- Mithen, S. (1996b). *The prehistory of the mind: A search for the origins of art, religion and science*. London: Thames & Hudson.
- Mithen, S. (2002). Human evolution and the cognitive basis of science. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 23-40). Cambridge: Cambridge University Press.

- Morton, A. (1980). *Frames of mind: Constraints on the common-sense conception of the mental*. Oxford: Clarendon Press.
- Nersessian, N. (1996). Child's play. *Philosophy of Science*, 63, 542-546.
- Nersessian, N. (1999). Model-based reasoning in conceptual change. In L. Magnani, N. Nersessian, & P. Thagard (Eds.), *Model-based reasoning in scientific discovery* (pp. 5-22). New York: Kluwer Academic/Plenum.
- Newcombe, N. (1998). Defining the 'radical middle'. *Human Development*, 41, 210-214.
- Newcombe, N., & Learmonth, A. (1999). Change and continuity in early spatial development: Claiming the "radical middle". *Infant Behavior & Development*, 22, 457-474.
- Nickles, T. (1981). What is a problem that we might solve it? *Synthese*, 47, 85-118.
- Nickles, T. (1990). Discovery. In R. C. Olby, G. N. Cantor, J. R. R. Christie, & M. J. S. Hodge (Eds.), *Companion to the history of modern science* (pp. 148-165). London: Routledge.
- Niiniluoto, I. (1999). Defending abduction. *Philosophy of Science*, 66 (Proceedings), S436-S451.
- Nowak, G., & Thagard, P. (1992). Copernicus, Ptolemy, and explanatory coherence. In R. N. Giere (Ed.), *Minnesota Studies in the Philosophy of Science: Vol. 15. Cognitive models of science* (pp. 274-309). Minneapolis: University of Minnesota Press.
- Oliver, A., Johnson, M. H., Karmiloff-Smith, A., & Pennington, B. (2000). Deviations in the emergence of representations: A neuroconstructivist framework for analyzing developmental disorders. *Developmental Science*, 3, 1-23.

- Peirce, C. S. (1931-1958). *Collected Papers*. Cambridge, MA: Harvard University Press.
- Peirce, C. S. (1982). *Writings of Charles S. Peirce. A chronological edition*. Bloomington: Indiana University Press.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge: MIT Press.
- Plunkett, K., Karmiloff-Smith, A., Bates, E., Elman, J. L., & Johnson, M. H. (1997). Connectionism and developmental psychology. *Journal of Child Psychology & Psychiatry*, 38, 53-80.
- Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.
- Popper, K. R. (1965). *Conjectures and refutations: The growth of scientific knowledge* (2nd ed.). New York: Basic Books.
- Popper, K. R. (1973). *Objective knowledge*. Oxford: Oxford University Press.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66, 211-227.
- Quine, W. V. O. (1963). *From a logical point of view*. New York: Harper & Rowe.
- Quine, W. V. O. (1969). *Ontological relativity and other essays*. New York: Columbia University Press.
- Quine, W. V. O., & Ullian, J. S. (1970). *The web of belief*. New York: Random House.
- Reichenbach, H. (1938). *Experience and prediction*. Chicago: University of Chicago Press.

- Rozeboom, W. W. (1972). Scientific inference: The myth and the reality. In S. R. Brown & D. J. Brenner (Eds.), *Science, psychology, and communication: Essays honoring William Stephenson* (pp. 95-118). New York: Columbia University Press.
- Rozeboom, W. W. (1990). Hypothetico-deductivism is a fraud. *American Psychologist*, 45, 555-556.
- Rozeboom, W. W. (1997). Good science is abductive, not hypothetico-deductive. In L. L. Harlow, S. A. Mulaik, & J. H. Steiger (Eds.), *What if there were no significance tests?* (pp. 335-391). Hillsdale, NJ: Erlbaum.
- Russell, J. (1992). The theory theory: So good they named it twice? *Cognitive Development*, 7, 485-519.
- Russell, J., Jarrold, C., & Potel, D. (1994). What makes strategic deception difficult for children – the deception or the strategy? *British Journal of Developmental Psychology*, 12, 301-314.
- Salmon, W. C. (1967). *The foundations of scientific inference*. Pittsburgh: University of Pittsburgh Press.
- Samarapungavan, A. (1992). Children's judgments in theory choice tasks: Scientific rationality in childhood. *Cognition*, 45, 1-32.
- Samarapungavan, A., & Wiers, R. W. (1997). Children's thoughts on the origin of the species: A study of explanatory coherence. *Cognitive Science*, 21, 147-177.
- Shelley, C. (1999a). Multiple analogies in archaeology. *Philosophy of Science*, 66, 579-605.
- Shelley, C. (1999b). Multiple analogies in evolutionary biology. *Studies in History and Philosophy of Science, Part C, Studies in History and Philosophy of Biological and Biomedical Sciences*, 30, 143-180.

- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.
- Simons, D. J., & Keil, F. C. (1995). An abstract to concrete shift in the development of biological thought: The insides story. *Cognition*, 56, 129-163.
- Slaughter, V., & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child Development*, 67, 2967-2988.
- Smith, C., Carey, S., & Wiser, M. (1985). On differentiation: A case study of the development of the concepts of size, weight, and density. *Cognition*, 21, 177-237.
- Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development*, 62, 753-766.
- Solomon, J. (1983). Thinking in two worlds of knowledge. In H. Helm & J. D. Novak (Eds.), *Proceedings of the International Seminar: Misconceptions in Science and Mathematics* (pp. 127-132). Ithaca, NY: Cornell University.
- Solomon, M. (1996). Commentary on Alison Gopnik's "The Scientist As Child". *Philosophy of Science*, 63, 547-551.
- Spelke, E. S. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 133-169). Hillsdale, NJ: Erlbaum.
- Spelke, E. S. (1994). Initial knowledge: Six suggestions. *Cognition*, 50, 431-445.
- Spencer, J., & Karmiloff-Smith, A. (1997). Are we misconstruing children or scientists? *Human Development*, 40, 52-54.
- Stich, S. (1990). *The fragmentation of reason: Preface to a pragmatic theory of cognitive evaluation*. Cambridge, MA: MIT Press.

- Stich, S., & Nichols, S. (1998). Theory Theory to the max. Critical notice of Alison Gopnik and Andrew N. Meltzoff, "Words, thoughts, and theories". *Mind & Language*, 13, 421-449.
- Thagard, P. (1978). The best explanation: Criteria for theory choice. *Journal of Philosophy*, 75, 76-92.
- Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press.
- Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, 12, 435-467.
- Thagard, P. (1992). *Conceptual revolutions*. Princeton: Princeton University Press.
- Thagard, P. (1999). *How scientists explain disease*. Princeton: Princeton University Press.
- Thagard, P. (2000). *Coherence in thought and action*. Cambridge, MA: MIT Press.
- Thagard, P. (2001). Foreword. In L. Magnani, *Abduction, reason, and science: Processes of discovery and explanation*. New York: Kluwer Academic/Plenum.
- Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial Intelligence*, 46, 259-310.
- Thagard, P. & Millgram, E. (1995). Inference to the best plan: A coherence theory of decision. In A. Ram & D. B. Leake (Eds.), *Goal-driven learning* (pp. 439-454). Cambridge, MA: MIT Press.
- Thagard, P., & Verbeurgt, K. (1998). Coherence as constraint satisfaction. *Cognitive Science*, 22, 1-24.

- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Toulmin, S. (1972). *Human understanding*. Princeton: Princeton University Press.
- Tukey, J. W. (1980). We need both exploratory and confirmatory. *American Statistician*, 34, 23-25.
- Vosniadou, S., & Brewer, W. F. (1987). Theories of knowledge restructuring in development. *Review of Educational Research*, 57, 51-67.
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, 535-585.
- Vosniadou, S., & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science*, 18, 123-183.
- Waddington, C. H. (1975). *The evolution of an evolutionist*. Edinburgh: Edinburgh University Press.
- Ward, T., & Haig, B. (1997). Abductive reasoning and clinical assessment. *Australian Psychologist*, 32, 93-100.
- Ward, T., Vertue, F. M., & Haig, B. D. (1999). Abductive method and clinical assessment in practice. *Behaviour Change*, 16, 49-63.
- Wellman, H. M. (1990). *The child's theory of mind*. Cambridge: MIT Press.
- Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, 43, 337-375.
- Wellman, H. M., & Gelman, S. A. (1998). Knowledge acquisition in foundational domains. In W. Damon (Series Ed.) & D. Kuhn & R. S. Siegler (Vol. Eds.),

Handbook of child psychology: Vol. 2. Cognition, perception, and language (5th ed., pp. 523-573). New York: Wiley.

Whewell, W. (1967). *The philosophy of the inductive sciences*. New York: Johnson Reprint Corp.

Wilson, R. A., & Keil, F. C. (2000). The shadows and shallows of explanation. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 87-114). Cambridge, MA: MIT Press.

Wiser, M. (1988). The differentiation of heat and temperature: History of science and the novice-expert shift. In S. Strauss (Ed.), *Ontogeny, phylogeny, and historical development* (pp. 28-48). Norwood, NJ: Ablex.

Woodward, J. (1989). Data and phenomena. *Synthese*, 79, 393-472.

Woodward, J. (2000). Data, phenomena, and reliability. *Philosophy of Science*, 67 (Proceedings), S163-S179.